lentin
rivaud

álgebra moderna

aguilar

A. LENTIN y J. RIVAUD

ÁLGEBRA MODERNA

La inclusión en los programas de los cursos universitarios iniciales de conceptos de álgebra relativamente modernos, plantea la necesidad urgente de disponer de un texto de introducción que sirva de guía a los alumnos que no tienen otros conocimientos matemáticos que los adquiridos en el bachillerato.

Los temas aquí tratados se consideran indispensables para quienes hayan de seguir estudios de física o matemáticas: conjuntos y funciones; grupos, anillos, cuerpos, números reales y complejos; espacios vectoriales, aplicaciones lineales, matrices; polinomios, fracciones racionales, formas lineales y cuadráticas, operadores lineales.

Los progresos realizados desde los comienzos de siglo obligan a renovar la enseñanza de la matemática, ya sea mediante nuevos conceptos que hagan posible ampliar el alcance del razonamiento, o bien actualizando este de manera que sean accesibles al principiante resultados considerados antes como difíciles.

Los autores, basados en su rica experiencia docente, hacen uso acertado de ambos métodos. La lectura del libro se hace sin esfuerzo y pone al alcance inmediato de la persona simplemente curiosa algunas de las más potentes técnicas que la matemática actual ofrece a las ciencias aplicadas. Verá al propio tiempo, expresados en un lenguaje más preciso y universal, conceptos que hace solo unos años parecían inconexos.

A. LENTIN
Maître de Recherches del C. N. R. S. (Francia)
y
J. RIVAUD
Doctor en Ciencias
Maître de Conférences de la Escuela Politécnica, París

ALGEBRA MODERNA

Versión española de
EMILIO MOTILVA YLARRI
Licenciado en Ciencias Exactas
Profesor de la Escuela Politécnica del Ejército

AGUILAR

La edición original de la presente obra, incorporado a este fondo editorial con el asesoramiento de DON LUIS BRAVO GALA, *ha sido publicada por la Casa Librairie Vuibert, de Paris, con el título siguiente:*

LECONS D'ALGÈBRE MODERNE

DEUXIÈME EDITION, 1964

SEGUNDA EDICION, 1967
Primera reimpresión, 1969
Segunda reimpresión, 1970

Depósito legal. M. 23952.-1970.

© Librairie Vuibert, 1964.

AGUILAR, S. A. DE EDICIONES, Juan Bravo, 38, Madrid (España), 1970.

Printed in Spain. Impreso en España por Gráficas Ume,S.A. Castillo Piñeiro, 8, Madrid.-1970.

PREFACIO

PREFACIO DE LA EDICION FRANCESA

El desarrollo de la matemática no es comparable a la formación de un montón de piedras, al que cada generación aporta la suya, sino que cada época reconstruye el edificio con arreglo a planos y cimientos nuevos. La nuestra, rica en nuevos materiales, edifica activamente, y urge por ello preparar a los jóvenes a vivir y orientarse en su futura casa.

Los numerosos maestros que, desde hace unos diez años, se consagran a resolver el delicado problema de la formación, comienzan a recoger hoy el fruto de sus esfuerzos: ciertas ideas que hasta ayer se consideraban como de «vanguardia» son ahora del dominio común de las nociones «modernas» que figuran en los programas de la enseñanza secundaria, recibiendo así el espaldarazo oficial.

Los estudiantes de hoy llegan, pues, a las clases preparatorias, de propedéutica y de licenciatura, con unos conocimientos y una disposición mental de los que carecían sus mayores.

Teniendo en cuenta estas consideraciones—y a pesar de que continuamos editando nuestra obra Elements d'Algèbre Moderne—*hemos publicado un libro algo más extenso que trata análogas cuestiones.*

La favorable acogida que han tenido estas Leçons d'Algèbre Moderne *nos obligó a preparar una segunda edición en la que se subraya aún más la diferencia que separa la nueva obra de los antiguos* Elements; *he aquí su contenido.*

La parte I trata de los conjuntos y estructuras fundamentales. La noción de aplicación se expone con cierta insistencia, y la hallaremos repetidas veces a lo largo de todo el libro como idea directriz. Del análisis combinatorio y del álgebra de Boole se dice solo aquello que es imposible ignorar. En el estudio de los grupos y anillos pueden dejarse para una segunda lectura las cuestiones de homomorfismos e ideales, a las que se dedican los finales de capítulo.

Termina esta parte un capítulo dedicado a la estructura de espacio vectorial, donde se encontrará todo aquello que se deduce de manera inmediata de los únicos conceptos de dependencia e independencia lineales.

La parte II construye los cuerpos de los números reales y de los números complejos.

En la parte III se expone la teoría, ya clásica, de los polinomios formales. Se ha prestado especial atención a su presentación pedagógica.

Muy interesante, la parte IV está dedicada al álgebra lineal. Sacrificando teorías elegantes, que solo tendrían interés para los algebristas puros, hemos pensado en aquellos lectores que han de utilizar el álgebra y que, en su mayoría, serán técnicos. Ahora bien: la casi totalidad de los problemas que en física y en mecánica se expresan mediante ecuaciones son problemas lineales, y es por ello indispensable lograr que los estudiantes adquieran lo que se ha llamado «reflejos lineales». En este orden de ideas, no nos ha importado desarrollar en muchas ocasiones (aplicaciones lineales, dualidad, aplicaciones bilineales o multilineales, etc.) razonamientos a veces muy parecidos y que hubieran podido condensarse más. Para conseguir la comprensión de los métodos de álgebra lineal, hemos recurrido al más antiguo de los procedimientos pedagógicos: la repetición.

Gracias a estas repeticiones, los diversos capítulos conservan una relativa independencia y el lector principiante no corre el peligro de perder definitivamente el hilo si tropieza en algún punto concreto; además, el libro puede leerse en dos niveles diferentes.

En el cálculo matricial hemos conservado un tono bastante concreto y se expone independientemente de la teoría más abstracta de los operadores. La ligazón se realiza más adelante, a propósito de la representación matricial de los operadores lineales.

Aunque la introducción de los determinantes está justificada por consideraciones teóricas relacionadas con las funciones multilineales y alternadas, el lector puede prescindir de esta motivación y partir de la definición obtenida, atribuyéndole el carácter de definición a priori.

El espacio dual se introduce «concretamente» como espacio de las formas lineales, y cabe dejar para un estudio ulterior la dualidad propiamente dicha.

Finalmente, si bien el último capítulo se inicia con la teoría general de las formas lineales o cuadráticas, se expone de manera independiente la teoría del espacio euclidiano que encuentra importantes aplicaciones en geometría.

En la segunda edición, se ha refundido totalmente este último capítulo. El espacio euclidiano puede identificarse con su dual: tal es la propiedad fundamental que se desprende al primer golpe y directamente. Según el orden que adopte el lector, esta propiedad servirá de aclaración o de introducción a la teoría general de la dualidad.

También se incluyen algunas nociones relacionadas con el espacio

unitario; además de su interés intrínseco, permiten dar significado geométrico a demostraciones que, sin ellas, parecerían demasiado artificiosas.

Numerosos lectores nos han honrado con sus acertados consejos. A todos ellos les damos aquí las gracias y permítasenos citar en particular al señor Brissaud, profesor del C.E.T. de Mâcon, a cuyas observaciones y acertadas críticas se debe más de una rectificación o mejora.

Finalmente, cumplimos con el agradable deber de expresar nuestra gratitud a la Librairie Vuibert por acceder a que se hayan hecho retoques tan importantes de una edición a otra.

INDICE GENERAL

INDICE GENERAL

PREFACIO DE LA EDICIÓN FRANCESA ... ... *Pág.* IX

PARTE I

CONJUNTOS. ESTRUCTURAS FUNDAMENTALES

CAP. I.—CONJUNTOS, APLICACIONES, RELACIONES BINARIAS 3

Conjuntos: 1. Noción de conjunto, *pág.* 3.—2. Subconjuntos, 4.—3. Conjunto producto, 5.—*Aplicaciones:* 4. Definición, 7.—5. Imagen de una parte de E. Aplicación suprayectiva, 8.—6. Imagen recíproca de una parte de F. Aplicación inyectiva, 8.—7. Aplicación biyectiva, 9.—*Relaciones binarias:* 8. Clase asociada a una propiedad. Cuantificadores, 11.—9. Relación binaria, 12.—10. Propiedades usuales de las relaciones binarias, 13.—*Relaciones de equivalencia:* 11. Relación de equivalencia, 15.—12. Clase de equivalencia, conjunto cociente, 16.—*Relaciones de orden:* 13. Relaciones de orden, 18.—14. Mayorante. Límite superior, 19.—Ejercicios sobre el capítulo I, 20.

CAP. II.—LEYES DE COMPOSICIÓN, ESTRUCTURAS, ÁLGEBRA DE BOOLE 22

Leyes de composición: 1. Definición, *pág.* 22.—2. Extensión de una ley de E a E×E, 23.—3. Ley estable respecto de una equivalencia, 24.—*Propiedades usuales de las leyes de composición:* 4. Leyes asociativas, 25.—5. Otras propiedades usuales, 27.—*Elemento neutro, elemento regular:* 6. Algunas propiedades posibles de los elementos, 29.—7. Elementos simétricos, 30.—8. La notación exponencial. 31.—*Composición de aplicaciones:* 9. Producto de dos aplicaciones, 33.—10. Aplicación de un conjunto en sí mismo, 33.—*Nociones de estructura y de isomorfismo:* 11. Estructura, 35.—12. Isomorfismo, 36.—13. Imagen homomorfa de un conjunto dotado de una ley, 38.—*Algebra de las partes de un referencial:* 14. Una estructura muy general, 40.—15. Algebra de Boole, 42.—Ejercicios sobre el capítulo II, 43.

CAP. III.—ENUMERACIONES DIVERSAS 45

Combinatoria: 1. Permutaciones, *pág.* 45.—2. Variaciones, 45.—3. Combinaciones, 46.—4. Permutaciones con repetición. 47.—5. Variaciones con repetición, 48.—6. Combinaciones con repetición, 48.—*Enumeración y álgebra de los conjuntos:* 7. Unión de dos conjuntos finitos, 50.—*Fórmulas sumatorias. Aplicaciones:* 8. Algunas fórmulas sumatorias, 51.—9. Un problema a título de ejemplo, 52.—Ejercicios sobre el capítulo III, 53.

CAP. IV.—ESTRUCTURA DE GRUPO 54

Los axiomas de la estructura de grupo: 1. Definición, *pág.* 54.—2. Otras propiedades elementales, 55.—3. Vocabulario y notaciones, 56.—4. Ejemplos de grupos. Grupos finitos, 57.—5. Subgrupos, 59.—*Grupos de sustituciones:* 6. Definición, 60.—7. Transposición, 62.—8. Paridad de una sustitución, 63. *Estudio de los subgrupos. Automorfismo. Homomorfismo de grupos:* 9. Partición estable de un grupo, 65.—10. Subgrupos conjugados. Automorfismos internos, 67.—11. Equivalencia regular, 69.—12. Homomorfismo para los grupos, 70.—13. Grupos monógenos. Grupos cíclicos, 71.—14. Compuesto de dos grupos, 73.—Ejercicios sobre el capítulo IV, 73.

CAP. V.—ESTRUCTURA DE ANILLO 75

Los axiomas de la estructura de anillo: 1. Definición, *pág.* 75.—2. Consecuencias de la existencia de un grupo, 76.—3. Otras propiedades, 77.—4. Anillo de integridad, 78.—5. Anillo ordenado, 79.—*El anillo Z de los enteros relativos:* 6. Introducción, 82.—7. Representación de los enteros por pares, 82.—8. Los enteros relativos, 83.—*El anillo de las clases residuales*

módulo n: 9. Congruencias, 86.—10. Clases residuales módulo *n*, 88.—11. Correspondencia entre Z y Z/*n*, 91.—*Relación de divisibilidad en los dominios de integridad:* 12. La relación de divisibilidad, 91.—13. Elementos asociados, 92.—*Homomorfismo de anillos:* 14. Imagen homomorfa de un anillo, 94.—Ejercicios sobre el capítulo V, 96.

CAP. VI.—ESTRUCTURA DE CUERPO 99

Los axiomas de la estructura de cuerpo: 1. Definición, *pág.* 98.—2. Reglas de cálculo, 99.—3. Ejemplos de cuerpos, 101.—4. Característica de un cuerpo. Subcuerpo, 103.—*El cuerpo Q de los números racionales:* 5. Introducción, 104.—6. Definición de números racionales, 105.—7. Operaciones, 105.—8. El cuerpo Q de los números racionales, 106.—9. Cuerpo ordenado, 108.—10. Consecuencias, 108.—11. Convergencia sobre un cuerpo ordenado, 110.—Ejercicios sobre el capítulo VI, 111.

CAP. VII.—ESTRUCTURA DE ESPACIO VECTORIAL 115

Axiomática de un espacio vectorial: 1. Introducción, *pág.* 115.—2. Definición de espacio vectorial, 115.—3. Propiedades elementales, 116.—4. Ejemplos de espacios vectoriales, 116.—5. Producto de espacios vectoriales, 118.—6. Combinaciones lineales, 118.—*Subespacios vectoriales:* 7. Definición, 121.—8. Ejemplos de subespacios, 121.—9. Intersección de subespacios, 122.—10. Suma de dos subespacios, 122.—11. Subespacio engendrado por un sistema de vectores, 123.—12. Subespacios suplementarios; proyecciones, 125.—*Dependencia e independencia lineal:* 13. Sistema ligado, 127.—14. Sistema libre, 128.—15. Teorema fundamental de la independencia, 130.—*Espacio vectorial de «n» dimensiones:* 16. Base de un espacio vectorial, 132.—17. Espacio de dimensión finita, 132.—18. Ejemplos, 133.—19. Isomorfismo, 134.—20. Automorfismo. Cambio de base, 135.—21. Dimensión de un subespacio, 137.—22. Intersección y suma de dos subespacios, 138.—*Estudio de un subespacio. Sistemas de ecuaciones lineales:* 23. Estudio de un subespacio de K^n, 139.—24. Determinación de una base, 140.—25. Aplicación a los sistemas lineales, 143.—Ejercicios sobre el capítulo VII, 146.

PARTE II

DOS CUERPOS FUNDAMENTALES

CAP. VIII.—EL CUERPO DE LOS NÚMEROS REALES 151

Construcción de los números reales: 1. Consideraciones de orden intuitivo, *pág.* 151.—2. Sucesiones regulares. Sucesiones equivalentes, 152.—3. Propiedades de las sucesiones regulares, 153.—4. Grupo aditivo de R, 154.—5. Grupo multiplicativo de R, 155.—6. El cuerpo de los números reales, 156.—*Propiedades fundamentales del cuerpo de los números reales:* 7. R es un cuerpo ordenado, 157.—8. R es un cuerpo arquimediano, 159.—9. R es un cuerpo completo, 160.—10. Principio de los intervalos, 161.—*Una propiedad fundamental de las funciones continuas:* 11. Definición, 163.—12. Teorema, 163.—13. Caso particular, 163.—14. Raíz *n*-ésima de un número real, 164.—Ejercicios sobre el capítulo VIII, 165.

CAP. IX.—ISOMORFISMO FUNDAMENTAL EN EL CUERPO DE LOS NÚMEROS REALES 168

Comparación entre el grupo aditivo de los elementos de R y el grupo multiplicativo de sus elementos positivos: 1. Propiedades del grupo aditivo, *pág.* 168.—2. Propiedades del grupo multiplicativo R_M^+, 168.—3. El isomorfismo fundamental, 169.—4. Funciones logarítmica y exponencial, 172.—*Extensión del teorema de isomorfismo:* 5. Caso de un grupo abstracto, 174.—6. Consideraciones intuitivas, 175.—7. Propiedades del grupo G, 175.—8. Teorema de isomorfismo, 177.—*Medida de longitudes sobre una recta:* 9. Extensión del teorema de isomorfismo, 178.—10. Propiedades de la recta, 178.—11. La medida de segmentos, 180.—Ejercicios sobre el capítulo IX, 181.

CAP. X.—EL CUERPO DE LOS COMPLEJOS 184

Construcción del cuerpo C: 1. Definición del producto, *pág.* 184.—2. Propiedades del producto, 185.—3 Estructura de cuerpo, 186.—4. Automorfis-

mo de C: complejos conjugados, 188.—5. Aplicaciones, 191.—*Representación trigonométrica de los números complejos:* 6. Imagen de un número complejo, 193.—7. Módulo y argumento, 194.—8. Aplicación al grupo multiplicativo, 195.—9. El plano de Argand-Cauchy, 198.—10. Raíces *n*-ésimas de un número complejo, 199.—*Nota sobre los cuaternios:* 11. Definiciones, 202. Ejercicios sobre el capítulo X, 205.

PARTE III

ALGEBRA DE LOS POLINOMIOS

CAP. XI.—POLINOMIOS DE UNA VARIABLE ... 211

Nociones generales: 1. La noción clásica de polinomio, *pág.* 211.—2. Polinomio formal, 211.—*Espacio vectorial de los polinomios formales sobre un cuerpo:* 3. Adición, 213.—4. Multiplicación por un escalar, 214.—5. Polinomios sobre un cuerpo K, 214.—*Anillo de los polinomios sobre un anillo:* 6. Producto de dos polinomios, 215.—7. Propiedades de la multiplicación, 217. 8. Anillo de los polinomios, 217.—9. Anillo de los polinomios sobre un cuerpo, 218.—10. Principio de especialización, 220.—*Clases de congruencia. División euclidiana:* 11. Clases de congruencia con relación a un polinomio, 222. 12. División euclidiana, 224.—13. Propiedades, 225.—14. División por $x - a$, 226.—15. Polinomios y funciones racionales enteras, 227.—16. División según las potencias crecientes, 229.—*Ideales de polinomios, M. C. D. y M. C. M.:* 17. Noción de ideal, 230.—18. M. C. D. Polinomios primos entre sí, 232.—19. Caso de dos polinomios, 233.—20. Mínimo común múltiplo, 236. *Nociones sobre eliminación:* 21. Polinomios no primos, 237.—Ejercicios sobre el capítulo XI, 240.

CAP. XII.—DESCOMPOSICIÓN EN FACTORES PRIMOS. FRACCIONES RACIONALES. 242

Factores primos. Teorema fundamental: 1. Polinomio irreducible sobre un cuerpo, *pág.* 242.—2. Teorema fundamental, 243.—3. Descomposición completa, 245.—*Extensión algebraica de un cuerpo:* 4. Anillo de las clases residuales, 246.—5. Extensión algebraica de un cuerpo, 247.—6. Ejemplo 1.º: el cuerpo de los complejos, 247.—7. Clases residuales módulo $x^2 + 2$, 249.—8. Cuerpo de ruptura de un polinomio, 250.—*Descomposición en los cuerpos C y R:* 9. Teorema de D'Alembert, 251.—10. Descomposición en C, 253. 11. Descomposición sobre R, 253.—12. Ejemplo 1.º: ecuación de tercer grado, 254.—13. Ejemplo 2.º: ecuación de cuarto grado, 255.—*Derivada de un polinomio. Raíces múltiples:* 14. Definición algebraica formal de la derivada, 258.—15. Derivadas sucesivas, 259.—16. Fórmula de Taylor, 259. 17. Raíces múltiples, 260.—*El cuerpo de las fracciones racionales:* 18. Cuerpo de las fracciones del anillo $K[x]$, 262.—19. Descomposición en elementos simples, 263.—20. Caso particular importante, 265.—Ejercicios sobre el capítulo XII, 268.

CAP. XIII.—POLINOMIOS DE VARIAS VARIABLES ... 271

Anillo $K[x, y]$: 1. Polinomios formales, *pág.* 271.—2. Espacio vectorial, 271. 3. Producto, 272.—4. Orden lexicográfico, 273.—*Polinomios simétricos:* 5. Automorfismos definidos por sustituciones sobre las variables, 274.—6. Teorema, 276.—7. Ejemplos, 278.—*Factorización única en* $K[x, y]$: 8. Anillo de factorización única, 280.—9. Caso del anillo $Z[x]$, 281.—10. Generalización, 283.—Ejercicios sobre el capítulo XIII, 284.—Ejercicios de recapitulación (Partes II y III), 286.

PARTE IV

ALGEBRA LINEAL

CAP. XIV.—LAS MATRICES ... 291

Nociones generales y notaciones: 1. Definición, *pág.* 291.—2. Subdivisión en cajas, 293.—3. Trasposición, 294.—*Espacio vectorial de las matrices de dimensiones dadas:* 4. Adición, 295.—5. Espacio vectorial de las matrices (m, n), 295.—6. Subespacios vectoriales, 297.—*Rango de una matriz. Aplicación a los sistemas lineales:* 7. Teorema fundamental, 298.—8. Matriz

regular, 301.—8. Sistema lineal. Interpretación vectorial, 301.—10. Condición de posibilidad, 302.—11. La condición de Kronecker, 303.—*Producto de matrices:* 12. Producto de dos matrices, 305.—13. Propiedades del producto, 308.—14. Ejemplos de fórmulas matriciales, 312.—15. Matrices elementales, 314.—16. Operaciones con cajas, 315.—*Anillo de las matrices cuadradas de orden n:* 17. Estructura de anillo, 317.—18. Subanillos, 320.—19. Inversión de una matriz cuadrada, 321.—20. Cambio de base en un espacio vectorial, 323.—21. Cálculo de la inversa, 329.—22. Aplicación al estudio de las matrices rectangulares, 329.—Ejercicios sobre el capítulo XIV. 331.

CAP. XV.—APLICACIONES LINEALES. ESPACIOS DUALES 334

Aplicaciones lineales: 1. Definiciones generales, *pág.* 334.—2. Ejemplos de aplicaciones lineales, 335.—3. Imagen de un sistema, 336.—4. Determinación de una aplicación lineal. Existencia, 337.—5. Imagen de un subespacio. Rango. Aplicación suprayectiva, 337.—6. Imagen recíproca de un subespacio de F. Núcleo, 338.—7. Ecuaciones lineales, 340.—*Composición de aplicaciones lineales:* 8. Espacio vectorial de las aplicaciones de E en F, 341.—9. Compuesta de dos aplicaciones lineales, 342.—*Aplicaciones entre espacios de dimensiones finitas:* 10. El espacio inicial es de *n* dimensiones, 343.—11. Biyección, 344.—*El espacio de las funciones lineales:* 12. Funciones lineales, 345.—13. Expresión analítica en una base. Existencia, 346.—14. Cambio de base, 347.—15. Operaciones con las funciones lineales, 348.—16. Dimensión del espacio F. Base, 349.—17. Dualidad, 350.—18. Subespacios conjugados 351.—19. Determinación práctica, 352.—*Espacios duales:* 20. Dualidad, 354.—21. Bases duales, 356.—22. Cambios de base, 358.—*Aplicaciones traspuestas:* 23. Traspuesta de una aplicación lineal, 360.—24. Traspuesta de una suma y de un producto, 361.—25. Núcleo y rango de la aplicación traspuesta, 362.

CAP. XVI.—APLICACIONES MULTILINEALES. DETERMINANTES 364

Aplicaciones multilineales: 1. Definición, *pág.* 364.—2. Expresión analítica en una base, 366.—*Aplicaciones antisimétricas:* 3. Transformada de una aplicación por una sustitución, 367.—4. Aplicación antisimétrica, 368.—5. Antisimetrización de una función *p* veces lineal, 370.—6. Función determinante en un espacio de *n* dimensiones, 371.—7. Consecuencias, 373.—8. Orientación de un espacio vectorial, 374.—*Determinante de una matriz:* 9. Definición, 375.—10. Determinantes de dos matrices traspuestas, 376.—11. Linealidad y antisimetría, 376.—12. Determinante de un producto de matrices, 378.—13. Desarrollo de un determinante por los elementos de una fila o columna, 380.—14. Aplicaciones, 381.—15. Cálculo de un determinante, 383.—*Aplicación a los sistemas lineales:* 16. Identidades preliminares, 386. 17. Inversión de una matriz, 386.—18. Sistema de Cramer, 388.—19. Ejemplo, 388.—20. Discusión de un sistema de *m* ecuaciones con *n* incógnitas, 389. *Producto exterior:* 21. Definición, 391.—22. Propiedades, 392.—23. Determinación en una base, 392.—Ejercicios sobre el capítulo XVI, 394.

CAP. XVII.—OPERADORES Y MATRICES 399

Representación matricial de las aplicaciones lineales: 1. Operadores y matrices, *pág.* 399.—2. Ejemplos, 400.—3. Isomorfismo entre matrices y operadores, 403.—Composición de aplicaciones y producto de matrices, 403.—5. Trasposición, 404.—*Endomorfismos de un espacio vectorial:* 6. Isomorfismo de anillos, 405.—7. Determinante de un endomorfismo, 406.—8. Grupo lineal, 407.—9. Matrices semejantes, 408.—10. Subespacio invariante, 410. *Diagonalización de un operador:* 11. Vector propio y valor propio, 411.—12. Obtención de los valores propios y de los vectores propios, 414.—13. Diagonalización: método práctico, 415.—14. Ecuación verificada por un operador diagonalizable, 417.—Ejercicios sobre el capítulo XVII, 421.

CAP XVIII.—FORMAS BILINEALES Y FORMAS CUADRÁTICAS 425

Aplicaciones bilineales: 1. Forma bilineal, *pág.* 425.—2. Aplicaciones asociadas a una función. Rango, 426.—*Aplicaciones bilineales y simétricas. Formas cuadráticas:* 3. Aplicaciones simétricas, 428.—4. Forma cuadrática asociada, 429.—5. Vectores conjugados, 431.—6. Diagonalización de una forma, 432.—7 Descomposición de una forma cuadrática en cuadrados, 432. *Formas cuadráticas sobre el cuerpo de los números reales:* 8. Caso del cuerpo R. Ley de Sylvester, 434.—9. Carácter de definición. Desigualdad de Schwartz, 436.—Ejercicios sobre el capítulo XVIII, 438.

CAP. XIX.—ESPACIO EUCLIDIANO. ESPACIO UNITARIO 439

Producto escalar: 1. Definición. Propiedad fundamental. *pág.* 439.—2. Forma fundamental, 440. —3. Ortogonalidad, 442.—4. Longitudes, 443.—5. Angulos, 443.—*Bases ortonormadas. Producto vectorial:* 6. Método de ortogonalización de Schmidt, 445.—7. Suplemento ortogonal, 447.—8. Cambio de base. Matrices ortogonales, 448.—9. Función determinante en un espacio euclidiano. Producto vectorial, 450.—*Operadores lineales: Traspuesto, operador ortogonal:* 10. Traspuesto de un operador, 452.—11. Transformación ortogonal. Grupo ortogonal, 454.—*Operadores y formas simétricas:* 12. Operador simétrico, 456.—13. Diagonalización de un operador simétrico, 458.—14. Función bilineal en un espacio euclidiano, 460.—15. Función bilineal y simétrica, 460.—16. Diagonalización simultánea de dos funciones, 461.—*Espacio unitario de «n» dimensiones:* 17. Función hermítica, 464. —18. Espacio unitario: producto escalar, 465 —19. Ortogonalidad. Base ortonormada, 467.—20. Cambio de base. Matriz unitaria, 468.—21. Asociado de un operador lineal. Grupo unitario, 469.—22. Operador hermítico, 470.—Ejercicios de recapitulación, 472.

INDICE ALFABÉTICO DE AUTORES Y MATERIAS 479

PARTE I

CONJUNTOS
ESTRUCTURAS FUNDAMENTALES

CAPITULO I

CONJUNTOS, APLICACIONES, RELACIONES BINARIAS

CONJUNTOS

1. Noción de conjunto.—La teoría abstracta de conjuntos linda con la filosofía y, analizando los principios mismos del conocimiento, promueve terribles problemas que no es cuestión de abordar aquí. Para alcanzar los fines técnicos que nos interesan es suficiente el punto de vista llamado «ingenuo» [1], en oposición al formalista o axiomático.

Convendremos en que un *conjunto* está constituido por objetos materiales, o por fenómenos o signos o entes abstractos, reunidos en virtud de una propiedad común.

El conjunto $A=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ de cifras arábigas contiene diez elementos; es *finito*. La sucesión natural de los números enteros es un ejemplo de conjunto infinito. Una recta del plano euclidiano define fuera de ella varios conjuntos infinitos de puntos: los semiplanos dotados o no de frontera y el propio plano privado de la recta. En lógica clásica, el conjunto de los valores lógicos que pueden tomar las proposiciones contiene dos elementos, verdadero y falso, denotados por 1 y 0, respectivamente.

Un conjunto es una entidad de naturaleza diferente a la de los elementos que lo componen; un conjunto de puntos no es un punto, incluso si contiene solo un punto.

Vocabulario y notaciones.—Si a es un elemento del conjunto E, escribimos $a \in E$, que se lee «a elemento de E» y también «a pertenece a E». La negación de la afirmación anterior se representa por $a \notin E$.

Es cómodo llamar *referencial* de una disciplina científica al conjunto del que esta disciplina trata primordialmente; el referencial de la entomología es el conjunto de los insectos.

---

[1] Esta «ingenuidad», relativa sin embargo, era la de Georg CANTOR (1845-1918), creador de la teoría.

Algunos conjuntos, particularmente importantes, se designan con letras determinadas. Entre ellos, recordemos:

N^*, conjunto de los enteros naturales $\{1, 2, 3, \ldots\}$;
N, conjunto de los enteros naturales completado con el cero;
Z, conjunto de los enteros relativos: números enteros positivos y negativos, incluido el cero $\{0, \pm 1, \pm 2, \ldots\}$;
Q, conjunto de los números racionales: fracciones positivas y negativas;
R, conjunto de los números reales;
C, conjunto de los números complejos, definidos más adelante.

NOTA.—En matemática elemental, cuando hablamos de *números*, nos referimos generalmente al conjunto R de los números reales:

$$-\frac{5}{2}, \quad -\sqrt{2}, \quad 0, \quad \sqrt[3]{5+\sqrt{3}}, \quad \log 45, \quad \operatorname{sen} 10$$

son números reales. La palabra *real* queda justificada por la correspondencia de nuestro espacio físico con la geometría: a todo punto de una recta (orientada, dotada de origen y de unidad de longitud) corresponde un número real, y viceversa.

2. Subconjuntos.—Si todos los elementos de un conjunto F pertenecen a un conjunto E, decimos que F está *incluido* en E, o que es un *subconjunto* de E, o también una *parte* de E; se escribe $F \subseteq E$. Según esta definición, el propio E debe considerarse como un subconjunto de sí mismo: es la inclusión en sentido *amplio*. Si nos consta que E contiene elementos extraños a F, decimos que F es un subconjunto *estricto* y dicha inclusión se escribe $F \subset E$.

La parte del conjunto E formada por elementos determinados $a, b, \ldots, l$ se simboliza por $\{a, b, \ldots, l\}$. Cabe, igualmente, designar una parte de un conjunto precisando las propiedades particulares P a las que satisface un elemento x de la misma:

$$\{x;\ x \in E;\ x \text{ verifica P}\}.$$

Así, p. ej., $\{x;\ x \in R;\ 0 \leqslant x \leqslant 1\}$ designa el conjunto de los números reales comprendidos entre 0 y 1.

Subconjuntos disjuntos.—Dos subconjuntos A y B de un mismo conjunto E se llaman *disjuntos* si carecen de elementos comunes.

Así, p. ej., en el conjunto N de los enteros, los subconjuntos de los números pares y de los números impares son disjuntos.

Complementario.—El subconjunto A', que contiene todos los elementos de E que no pertenecen a A es, evidentemente, disjunto de A: se denomina complementario de A; naturalmente, $(A')' = A$. Para evitar excepciones a esta definición es necesario introducir un conjunto convencional, el complementario de E; es el conjunto vacío, que denotaremos con el símbolo ϕ.

Conjunto de las partes.—Llamamos *conjunto* de las partes de un conjunto E, y lo designamos por $\mathfrak{P}(E)$, al conjunto cuyos elementos son las *partes* de E. Se tiene $\phi \in \mathfrak{P}(E)$, $E \in \mathfrak{P}(E)$ y, cualquiera que sea $x \in E$, $\{x\} \in \mathfrak{P}(E)$.

Intersección.—Sean A, B, C, ..., partes bien determinadas de un conjunto E. Se llama *intersección* de una familia de partes, A, B, C, por ejemplo, el subconjunto formado por los elementos de E que pertenecen a cada una de las partes consideradas. Esta intersección se designa con la notación $A \cap B \cap C$. En particular, si dos de los subconjuntos considerados son disjuntos, la intersección queda reducida a la parte vacía.

Reunión.—Del mismo modo, se llama *reunión* de las partes A, B, C, al subconjunto formado por los elementos de E que pertenecen al menos a una de dichas partes. Escribimos $A \cup B \cup C$ para representar la reunión de A, B, C.

Partición.—Se realiza una *partición* de un conjunto E clasificando los elementos de dicho conjunto en subconjuntos disjuntos E_1, E_2, ..., tales que todo elemento de E esté clasificado.

Los enteros naturales dan lugar a la partición en números pares y en números impares; los polígonos planos, en polígonos convexos, cóncavos, no estrellados, estrellados.

3. Conjunto producto.—Sean dos conjuntos

$$E = \{a, b, \ldots, x, y, \ldots\} \qquad \text{y} \qquad F = \{\alpha, \beta, \ldots, \xi, \ldots\}.$$

Los pares ordenados del tipo (x, ξ) con $x \in E$ y $\xi \in F$ forman un tercer conjunto que simbolizamos por $E \times F$ y se llama producto de E por F. Se considera que el par (x, ξ) es diferente del par (ξ, x), que pertenece al producto $F \times E$. Si E y F son dos conjuntos disjuntos no hay confusión posible; pero si E y F son idénticos, los pares (x, y) e (y, x) son elementos distintos de $E \times E$. La primera coordenada, x, y la segunda, y, del par (x, y) se distinguen por el orden de su escritura;

los elementos de la forma (x, x) se llaman diagonales. La noción se generaliza a un número finito de conjuntos $E_1, E_2, \ldots, E_n$. El conjunto producto $E_1 \times E_2 \times \ldots \times E_n$ se compone de todas las sucesiones ordenadas formadas por n elementos $(x_1, x_2, \ldots, x_n)$ con $x_1 \in E_1, \ldots, x_n \in E_n$. El producto $E \times E \times \ldots \times E$ se designa por E^n.

NOTA. Sean tres conjuntos E_1, E_2, E_3. Consideremos los conjuntos

$$(E_1 \times E_2) \times E_3, \qquad E_1 \times E_2 \times E_3 \qquad y \qquad E_1 \times (E_2 \times E_3);$$

sus elementos difieren únicamente en detalles de escritura:

$$((x_1, x_2), x_3), \qquad (x_1, x_2, x_3), \qquad (x_1, (x_2, x_3)),$$

que podemos despreciar si solo tomamos en consideración los elementos utilizados de cada uno de los conjuntos. Por consiguiente, se podría escribir:

$$E_1 \times E_2 \times E_3 = (E_1 \times E_2) \times E_3 = E_1 \times (E_2 \times E_3)$$

y la última igualdad corresponde a la asociatividad (pág. 26) de la multiplicación de conjuntos.

EJEMPLO. En geometría analítica fijamos un punto M del plano por medio de dos números reales: x es la abscisa, e y, la ordenada. En estas condiciones el plano aparece como el conjunto de los pares ordenados (x, y) con $x \in R$ e $y \in R$, designando con R al conjunto de los números reales; a todo punto corresponde un par, y viceversa. El plano, por consiguiente, es el conjunto $R \times R = R^2$. De la misma manera el espacio de la geometría se simboliza por R^3.

EJERCICIOS:

1. E, conjunto finito, comprende p elementos; determínese el número de las partes de dicho conjunto, comprendidas ϕ y E.

Resp.: 2^p.

2. Sea E el conjunto de puntos de un plano P; dibujamos en P dos circunferencias concéntricas γ_1 y γ_2. Examínense las diferentes particiones posibles de E.

3. Si S y Σ son finitos, determínese el número de elementos de $S \times \Sigma$.

4. Siendo E el conjunto de los números x tales que $0 \leqslant x \leqslant 1$, demuéstrese que $E \times E$ puede representarse como conjunto de los puntos de un cuadrado de lado 1, construído sobre dos ejes rectangulares Ox y Oy. Sitúense los elementos diagonales.

5. Sean A y B ($A \subset B$) dos partes de E. Compárense los subconjuntos complementarios A' y B'.

APLICACIONES

4. Definición.—Dados dos conjuntos E y F, definimos una *aplicación* de E en F dando un criterio que permita hacer corresponder a todo elemento de E un elemento bien determinado de F. Esta regla se considera como un operador, que designaremos con una letra f, T, ... Si $a \in E$, $f(a)$ designa el transformado de a y representa, por consiguiente, un elemento de F. Se escribe

$$a \xrightarrow{f} f(a)$$

y se dice que a se aplica sobre $f(a)$. Es importante distinguir claramente los tres elementos que intervienen en una aplicación:

$$a \in E; \qquad f, \text{ el operador}; \qquad f(a) \in F.$$

Con frecuencia es cómodo designar con otra letra, b, al elemento de F denotado por $f(a)$. En tal caso se escribe $a \longrightarrow b = f(a)$; pero esto es un convenio de escritura que interesa separar de la noción de aplicación.

El elemento $f(a)$ se llama *imagen* de a; inversamente, a es un antecedente de $f(a)$; E se llama conjunto inicial, y F, conjunto final de la aplicación.

Aplicación y función.—La noción de función, ya conocida del lector, no difiere en principio del concepto de aplicación. Aquella se refiere al caso de ser E y F conjuntos de números. Adoptamos también algunas variantes de terminología: a es la variable (designada generalmente por x), E el dominio de definición y F el dominio de valores.

Conviene evitar una confusión demasiado frecuente entre f, que designa la función (o el operador), y $f(x)$, que es su valor para el elemento x. Corrientemente se dice la función sen x, mientras que debería decirse la «función seno». La notación de las aplicaciones $x \longrightarrow \text{sen } x$ evita dicha confusión. Se puede decir también: la función f definida por $f(x) = \text{sen } x$.

Transformaciones geométricas.— Las transformaciones geométricas del plano o del espacio proporcionan numerosos ejemplos de aplicaciones. Los conjuntos E y F se confunden muchas veces y sus elementos son puntos. En la transformación T, $M' = T(M)$ representa el transformado del punto M.

Caso particular.— Con frecuencia encontraremos aplicaciones de un conjunto E en sí mismo; es decir, $F = E$. En estas condiciones existe

una aplicación particular tal que, a todo $a \in E$, asocia el mismo elemento a. Es la *aplicación idéntica,* generalmente designada por I.

Restricción de una aplicación.—Sea E_1 un subconjunto cualquiera de E. Toda aplicación de E en F determina una aplicación de E_1 en F, en la que la imagen de un elemento de E_1 queda definida en virtud de pertenecer a E. Esta aplicación se llama *restricción* de f a E_1.

Gráfica de una aplicación.— Llamamos así al subconjunto del conjunto producto $E \times F$, constituido por los pares $[a, f(a)]$ asociados en la aplicación. Así, p. ej., en la teoría ordinaria de funciones, la gráfica de la aplicación $x \longrightarrow f(x)$ está constituida por el conjunto de puntos (x, y) del plano, tales que $y = f(x)$.

5. Imagen de una parte de E. Aplicación suprayectiva.—Sea f una aplicación de E en F, y E_1, un subconjunto cualquiera de E. En la aplicación f, el conjunto de las imágenes de los elementos de E_1 constituye un subconjunto de F, que escribimos con la notación $f(E_1)$ y que se llama *imagen* del subconjunto E_1. Si E_1 es vacío establecemos, además, que $f(\phi) = \phi$. En general, tendremos $f(E_1) \subseteq F$.

Casos particulares.—1.° Si E_1 se reduce a un elemento, su imagen queda igualmente reducida a un punto, de acuerdo con la definición de aplicación.

2.° Si $E_1 = E$, el subconjunto obtenido, $f(E)$, se llama *subconjunto imagen* de la aplicación. En cualquier caso se tiene que $f(E) \subseteq F$.

Aplicación suprayectiva.—Si se verifica que $f(E) = F$, de modo que el conjunto de las imágenes recubre totalmente a F, la aplicación se llama *suprayectiva*; decimos que es una *suprayección*. Se ha empleado igualmente la expresión: aplicación de E *sobre* F, reemplazando la preposición *sobre* a la preposición *en* utilizada para el caso general.

Por definición de $f(E)$, la aplicación de E en $f(E)$ es, evidentemente, suprayectiva.

6. Imagen recíproca de una parte de F. Aplicación inyectiva.— Consideremos ahora una parte F_1 de F. Llamamos *imagen recíproca* de F_1 por f al subconjunto de todos los $a \in E$ tales que $f(a) \in F_1$. Denotaremos por $f^{-1}(F_1)$ esta parte de E, conviniendo además que $f^{-1}(\phi) = \phi$.

Advirtamos que, si f no es suprayectiva, podemos tener $f^{-1}(F_1) = \phi$

incluso siendo $F_1 \neq \phi$. Para ello es suficiente que F_1 y el subconjunto imagen sean disjuntos.

EJEMPLO.—Proyectemos el espacio euclidiano E sobre uno de sus planos F, paralelamente a una recta Δ que corte a F. Un punto m de F admite como imagen recíproca todos los puntos de E situados en la paralela a Δ trazada por m.

Aplicación inyectiva.—Decimos que una aplicación de E en F es *inyectiva* (o también que es una *inyección*) cuando dos elementos distintos de E tienen por imágenes dos elementos diferentes de F. Por consiguiente, todo elemento de F puede tener, como máximo, un antecedente; es decir, la imagen recíproca de todo punto será vacía o quedará reducida a un punto.

7. Aplicación biyectiva.—Una aplicación f de E en F que sea a la vez suprayectiva e inyectiva se llama *biyectiva*. Decimos también que es una *biyección*. En tal aplicación, todo $b \in F$ admite un antecedente y uno solo en E, lo que define una aplicación de F en E. Esta nueva aplicación es suprayectiva, ya que todo elemento a de E es antecedente del elemento $f(a)$ de F; es también inyectiva, ya que dos elementos diferentes de F no pueden tener el mismo antecedente en E. Hemos obtenido, pues, una biyección de F en E, llamada *biyección recíproca* de f (o también función inversa) y que denotaremos por f^{-1}. En este caso podremos escribir las dos igualdades equivalentes

$$b = f(a) \qquad \text{o} \qquad a = f^{-1}(b).$$

Dos conjuntos entre los cuales es posible establecer una biyección se llaman *equipotentes*; se dirá también que tienen la misma potencia.

NOTAS. 1.ª Consideremos una aplicación inyectiva en la que $f(E)$ es el subconjunto imagen; la aplicación de E sobre $f(E)$ es biyectiva, pero, en general, no es una biyección de E sobre F.

2.ª En el parágrafo precedente hemos empleado el símbolo $f^{-1}(F_1)$ para designar la imagen recíproca de una parte de F y, en este caso, f^{-1} no designa una aplicación de F en E. Hagamos constar, sin embargo, que si se trata de una biyección la imagen recíproca de un subconjunto reducido a un elemento, $f^{-1}(\{b\})$, coincide con el transformado $f^{-1}(b)$ en la biyección recíproca. No hay, por consiguiente, contradicción en los dos empleos del signo f^{-1}.

3.ª A veces se dice de una aplicación que es *biunívoca*. Si no se añade nada más, esto significa que, por tener todo elemento de E una imagen, toda *imagen* tiene solo un antecedente: por consiguiente, la transformación es inyectiva. Pero si precisamos que la transformación es biunívoca entre E y F, debemos entender que, al tener todo elemento de F un antecedente y uno

solo, existe una biyección de E sobre F con su biyección recíproca de F sobre E.

EJERCICIOS:

1. Determínense los coeficientes p y q:

a) para que p y q sean *las* raíces de la ecuación $x^2+px+q=0$;

b) para que p y q sean cada una raíz de $x^2+px+q=0$.

Explíquese la diferencia de los resultados obtenidos analizando la aplicación del conjunto $\{p, q\}$ en el conjunto $\{x', x''\}$.

2. Escríbanse todas las aplicaciones del conjunto $\{a, b, c\}$ en sí mismo. Para representar una aplicación se escribirán debajo de $\{a, b, c\}$ sus imágenes respectivas; p. ej. $\begin{pmatrix} a & b & c \\ c & a & a \end{pmatrix}$ aplica a sobre c, b sobre a y c sobre a.

3. Dadas dos rectas Δ y Δ' y un punto O en un mismo plano, sea M un punto de Δ; OM corta generalmente a Δ' en un punto M'. Precísense los convenios que debemos establecer para obtener una aplicación biunívoca.

4. Imagen de $\{x;\ 2\leqslant x\leqslant 3\}$ por $x \longrightarrow y=\dfrac{x+2}{x-1}$. Estúdiese la aplicación inversa. Las mismas cuestiones para $\{x;\ 0\leqslant x\leqslant 3\}$.

5. Demuéstrese que, para un conjunto finito E, toda aplicación de E en E que sea suprayectiva es también biyectiva.

6. En el conjunto Z de los enteros relativos definimos la transformación:

$$n \in Z, \qquad f(n)=\frac{n}{2} \quad \text{si } n \text{ par}, \qquad f(n)=0 \quad \text{si } n \text{ impar}.$$

Demuéstrese que dicha transformación es suprayectiva pero no biyectiva. (La propiedad anterior no es, por consiguiente, cierta para un conjunto infinito.)

7. Sea f una aplicación de E en F. A todo $A \in \mathfrak{P}(F)$ asociamos $f(A) \in \mathfrak{P}(F)$.

a) ¿Qué podemos afirmar de $f(A)$ y $f(B)$ si $A \subset B$?

b) Compruébese la fórmula $f(A \cup B)=f(A) \cup f(B)$.

8. Con las mismas notaciones, demuéstrese que, en general, podemos escribir

$$f(A \cap B) \subset f(A) \cap f(B).$$

Si la aplicación f es inyectiva pruébese que el signo $\subset$ puede reemplazarse por el signo $=$.

9. En una biyección de E sobre F demuéstrese que

$$f(A')=[f(A)]',$$

donde $'$ designa el complementario.

10. En $\mathfrak{P}(E)$ definimos una aplicación $A \longrightarrow A'$. Pruébese que se trata de una biyección involutiva; es decir, idéntica a la biyección recíproca.

RELACIONES BINARIAS

8. Clase asociada a una propiedad. Cuantificadores.—El estudio de un conjunto dado E conduce muchas veces a particularizar ciertos subconjuntos cuyos elementos verifican una o varias propiedades no pertenecientes a los otros elementos de E. Sin duda, el lector podrá imaginar numerosos ejemplos de tal situación. Basta que la propiedad considerada, P, divida sin ambigüedad al conjunto E en dos clases complementarias A y A', tales que, para todo $x \in A$, se verifique P y que para todo $x \in A'$, no se verifique.

Consideremos ahora dos propiedades P y Q relativas a los elementos de un mismo conjunto E. Con frecuencia se encuentran proposiciones (teoremas) de la forma siguiente: *todo* $x \in E$ que verifica P (hipótesis) verifica también Q (conclusión). Esto significa que el subconjunto de los elementos que verifican P está incluido en el subconjunto de los elementos que verifican Q. Decimos que P implica Q y escribimos:

$$P \Longrightarrow Q. \qquad [1]$$

Así, p. ej., en el conjunto de los números enteros, si P es la propiedad de ser divisible por 6, y Q la propiedad de ser divisible por 3, sabemos que $P \Longrightarrow Q$.

El *recíproco* del teorema correspondiente a [1] se expresa por la implicación $Q \Longrightarrow P$. Para que esta última sea verdad es necesario que el subconjunto de los $x \in E$ que verifican Q esté incluido en el subconjunto correspondiente a P.

Si un teorema y su recíproco son simultáneamente ciertos, las propiedades P y Q se llaman *equivalentes* (los dos subconjuntos se confunden), lo que escribimos mediante una doble implicación:

$$P \Longleftrightarrow Q,$$

que se lee «se verifica P *si, y solo si,* Q se verifica», o también «para que P sea cierto es condición necesaria y suficiente que Q lo sea».

Cuantificadores.—Consideremos un conjunto E y una determinada propiedad P. Cabe plantear las siguientes cuestiones:

a) ¿Existen elementos de E que posean dicha propiedad?

b) En caso afirmativo, ¿la satisfacen todos?

Para responder a estas preguntas introducimos dos símbolos llamados *cuantificadores*. Estos son:

a) *Cuantificador existencial*, que se escribe $\exists$ y significa: existe al menos un elemento de E que tiene la propiedad P; p. ej., la escritura

$$(\exists x)\,(x \in \mathrm{R}) : 4x^2 - 5 = 0$$

significa que existe al menos un número real tal que $4x^2 - 5 = 0$.

b) *Cuantificador universal*, que se escribe $\forall$ y significa que todo elemento de E verifica P. Así, p. ej.,

$$(\forall x)\,(x \in \mathrm{R}) : (x+1)^2 = x^2 + 2x + 1,$$

significa que todo número real verifica la identidad escrita.

9. Relación binaria.—Con frecuencia una propiedad P concierne a dos elementos x e y pertenecientes, respectivamente, a dos conjuntos dados, E y F. Reducimos este caso al anterior considerando el conjunto producto $\mathrm{E} \times \mathrm{F}$; los pares (x, y) que satisfacen la propiedad P forman un subconjunto de $\mathrm{E} \times \mathrm{F}$. Diremos también que P define sobre los conjuntos E y F una relación binaria $\mathcal{R}$, y escribimos:

$x \mathcal{R} y$, cuando (x, y) satisface a P, $\overline{x \mathcal{R} y}$, si (x, y) no satisface a P.

Si F coincide con E, $\mathcal{R}$ se llama relación binaria en E. En adelante nos limitaremos a este caso.

Sobre un conjunto cualquiera, una relación binaria particularmente importante es la igualdad simbolizada por $=$. El empleo de este signo no es, o al menos no lo ha sido siempre, todo lo preciso que sería de desear. Para nosotros, *la igualdad se confunde con la identidad: $a = b$* significa que el elemento representado por a es el mismo que el representado por b. Tomando como referencial el conjunto de los reyes de Francia, escribiremos:

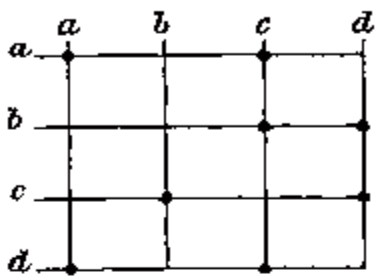
Luis XIV = El Rey Sol,
Luis IX = San Luis.

Se observará que esta definición va contra la práctica escolar que trata de «casos de igualdad» de triángulos no coincidentes. La unión o acuerdo entre el lenguaje clásico y el moderno se hará mediante la noción de clases de equivalencia. (Véase pág. 15.)

Hemos citado las relaciones geométricas de paralelismo y ortogonalidad definidas sobre un conjunto de rectas. Sobre el mismo conjunto citemos, además, la relación de incidencia $D_1 \nparallel D_2$ [D_1 corta a D_2].

En aritmética, aparte de las relaciones $<$, $\leqslant$, tenemos la relación $a \mid b$ [a divide a b], y relaciones del tipo $a \equiv b$ (mód. 3), que significa que a y b difieren en un múltiplo de 3 y que se lee «a congruente con b módulo 3», etc.

Si E es finito, podemos representar E×E y marcar los pares que pertenecen al subconjunto V en el que se verifica $\mathcal{R}$. He aquí las representaciones más frecuentes:



1.ª Representación por un gráfico o grafo:
· si verifica.

$\mathcal{R}$	a	b	c	d
a	1	0	1	0
b	0	0	1	1
c	0	1	0	1
d	1	0	1	0

2.ª Representación por una matriz:
1 si verifica; 0 no verifica.

EJERCICIOS:

1. $E=\{x;\ 0\leqslant x\leqslant 1\}$; E×E vendrá representado en un sistema de ejes cartesianos rectangulares por los puntos de un cuadrado de lado 1. Dibújese el conjunto V en el que se verifica $\mathcal{R}$, para las relaciones siguientes:

a) $y=x$; *d)* $x^2+y^2\leqslant 1$;
b) $y\leqslant x$; *e)* $y<x^2$.
c) $y<x$;

2. Una función f, $x\to y=f(x)$ que toma sus argumentos y sus valores en E, permite definir entre los pares (a, b) una relación verdadera si $b=f(a)$, falsa si $b\neq f(a)$. ¿Con qué condición una relación $\mathcal{R}$ permite definir una función? Examínense todos los casos del ejercicio anterior.

3. Escribiremos $A\mathcal{J}d$ si el punto A está sobre la recta d. Matriz de esta relación para los vértices y lados de un cuadrilátero completo.

4. La misma cuestión para dos triángulos homológicos t (teorema de Desargues).

10. Propiedades usuales de las relaciones binarias.—Las relaciones binarias más frecuentes poseen una o varias de las propiedades siguientes, que preferimos estudiar primero por separado.

Reflexividad.—Una relación $\mathcal{R}$ se llama *reflexiva* si, y solo si, el dominio en el que se verifica contiene la diagonal de E×E; en otras palabras, cualquiera que sea x se tiene $x\mathcal{R}x$.

La igualdad es reflexiva, la semejanza de triángulos también lo es. En el conjunto de los humanos muertos o vivos, la relación «x es hermano de y» no es reflexiva, ya que nadie es su propio hermano.

En un primer análisis, el paralelismo de rectas no es una relación reflexiva, pero modificando la definición de paralelismo diciendo que dos rectas son paralelas si son:

$$\begin{cases} \text{o coplanarias y sin punto común,} \\ \text{o confundidas} \end{cases}$$

resulta que el paralelismo es una relación reflexiva.

Simetría.—Una relación binaria $\mathcal{R}$ es *simétrica* si, y solo si, $a\mathcal{R}b$ implica $b\mathcal{R}a$.

La igualdad es simétrica, el paralelismo y la ortogonalidad de rectas lo son también. En el conjunto de los humanos, la relación «x es hermano o hermana de y» es simétrica, pero la relación «a es un ascendiente de b» no lo es.

El gráfico de una relación simétrica es simétrico con relación a la diagonal.

Antisimetría.—Una relación binaria $\mathcal{R}$ es *antisimétrica* si, y solo si, $a\mathcal{R}b$ y $b\mathcal{R}a$ implican $a=b$.

Por consiguiente, resulta que si $a \neq b$ y si $a\mathcal{R}b$, no puede verificarse la relación $b\mathcal{R}a$.

No es necesario decir que la antisimetría excluye la simetría, pero es compatible con la reflexividad, como lo demuestra la relación $x \leqslant y$ sobre el conjunto de los enteros. La relación «a es un ascendiente de b» es manifiestamente antisimétrica; la relación $\subseteq$, definida entre las partes de un referencial, lo es también.

Transitividad.—Una relación binaria $\mathcal{R}$ es *transitiva* si $a\mathcal{R}b$ y $b\mathcal{R}c$ implican $a\mathcal{R}c$.

La igualdad es transitiva, el paralelismo de rectas también, pero la ortogonalidad e incidencia no lo son. La relación $x \leqslant y$ y la relación «a es un ascendiente de b» son transitivas, así como la $\subseteq$.

EJERCICIOS:

1. Representada $\mathcal{R}$ por un grafo Γ, estúdiese Γ en los casos de reflexividad, simetría, transitividad.

2. Examínense desde el punto de vista de la reflexividad, simetría, antisimetría y transitividad las relaciones siguientes:

A) En el conjunto de los puntos del espacio:

a) M está del mismo lado que M' con relación a un plano dado Π;

b) M es conjugado de M' con relación a una esfera S;

c) M dista menos de 1 metro de M'.

B) En el conjunto de las circunferencias de un plano:

a) γ es ortogonal a γ';

b) γ es seudoortogonal a γ';
c) γ es concéntrica con γ'.

C) En el conjunto de los enteros relativos (positivos, negativos o nulos):
a) $x \equiv y$ (mód. 3);
b) $x^2+x=y^2+y$;
c) existe un entero K tal que $y=Kx$.

D) En el conjunto de palabras de nuestro idioma:
a) y puede ser sustituida por x en toda frase sin alterar la corrección (puramente) gramatical;
b) x concuerda en número con y.
c) x concuerda en género y número con y;
d) x es derivada de y.

RELACIONES DE EQUIVALENCIA

11. Relación de equivalencia.—Una relación binaria δ es de equivalencia si es reflexiva, simétrica y transitiva.

Esto significa que para todo a, tenemos $a\delta a$, que $a\delta b$ implica $b\delta a$, y que $a\delta b$ junto con $b\delta c$ implican $a\delta c$

La igualdad, el paralelismo, la equipolencia de vectores y la semejanza de triángulos son relaciones de equivalencia.

Para subrayar que δ es una relación de equivalencia empleamos a veces la notación específica $a \equiv b$ (mód. δ), que se lee «a congruente con b *módulo* δ», y también «a equivalente a b módulo δ». Si no hay confusión posible se dirá «a equivalente a b».

Equivalencia e igualdad.—Desde un punto de vista formal, toda relación de equivalencia δ generaliza, pues, las propiedades de la igualdad. Recíprocamente, la igualdad aparece con frecuencia como «contracción» de una equivalencia. Ya que no son *idénticos*, los símbolos 2/3, 4/6, 6/9 no son *iguales*, pero sí lo son los números racionales que representan, puesto que los símbolos son equivalentes según se desprende de la conocida regla «producto de medios igual a producto de extremos».

Igualmente, la equipolencia de los vectores ligados conduce a la gualdad de los vectores libres, y la congruencia de triángulos (la palabra fue introducida en geometría por D. Hilbert) conduce a lo que podríamos llamar igualdad de triángulos «libres».

Vamos a examinar el proceso mediante el cual, a partir de un

conjunto E dotado de una equivalencia $\mathcal{E}$, se construye un nuevo conjunto sobre el que $\mathcal{E}$ induce una relación de igualdad.

Encontramos también la noción de equivalencia en otras ciencias distintas de las matemáticas.

A título de ejemplo, digamos que la *fonología* de una lengua considera como equivalentes dos sonidos que el *fonético* reconoce como diferentes, con tal que estos dos sonidos no se hallen jamás en oposición y que la sustitución de uno por el otro sea indiferente en la comunicación de dos personas que hablan. La clase de equivalencia es un *fonema*.

12. Clase de equivalencia, conjunto cociente.—Sea E un conjunto dotado de una relación de equivalencia $\mathcal{E}$. Si $a \in E$, llamaremos *clase de equivalencia a* y la designaremos con $C(a)$ al conjunto finito o infinito de elementos de E equivalentes al a.

LEMA 1.—*Dos elementos cualesquiera de una clase de equivalencia son equivalentes entre sí.*

$a_1 \in C(a)$ y $a_2 \in C(a)$ implican $a \mathcal{E} a_1$ y $a \mathcal{E} a_2$; la simetría da $a_2 \mathcal{E} a$ y la transitividad $a_2 \mathcal{E} a_1$.

LEMA 2.—*Dos clases que tengan un elemento común son idénticas.*

Hipótesis: $a \in C(b)$ y $a \in C(c)$. Según el lema 1, $a \in C(b)$ implica $C(b) = C(a)$; igualmente $C(c) = C(a)$ y, por consiguiente, $C(b) = C(c)$.

De estos lemas resulta inmediatamente el teorema siguiente:

TEOREMA.—*Una relación de equivalencia* $\mathcal{E}$, *definida sobre un conjunto* E, *determina una partición de* E *en clases de equivalencia.*

Definición.—El conjunto de las clases de equivalencia determinado sobre E por $\mathcal{E}$ se llama conjunto cociente de E por $\mathcal{E}$; lo designaremos con el símbolo $E/\mathcal{E}$.

La equivalencia sobre E, $a \equiv b$ (mód. $\mathcal{E}$), induce la igualdad $C(a) = C(b)$ en $E/\mathcal{E}$.

Un ejemplo sencillo nos lo proporciona la equipolencia, que es una relación de equivalencia definida sobre el conjunto de los vectores. Un elemento del conjunto cociente está constituido por la clase de todos los vectores equipolentes a un vector dado: es un *vector libre*.

Equivalencia asociada a una aplicación.— Sean E y E' dos conjuntos y f una aplicación de E en E'. Sobre E establecemos la relación $a \mathcal{F} b$ si, y solo si, a y b tienen la misma imagen por f. Es inmediato ver que $\mathcal{F}$ es una relación de equivalencia. Si $f(E)$ es el conjunto de las imágenes,

en estas condiciones existe una aplicación $\bar{f}$ biunívoca de $E/\mathcal{F}$ sobre $f(E)$.

EJERCICIOS:

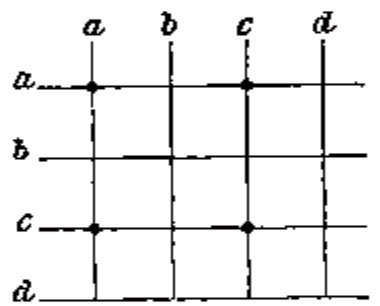
1. Demuéstrese que toda partición de un conjunto E define sobre E una relación de equivalencia (recíproco del teorema de la pág. 16).

2. Siendo θ la medida en radianes de un ángulo orientado, y ρ, x e y números relativos, aplicamos el conjunto Π de pares ordenados (ρ, θ) sobre el conjunto P de pares ordenados (x, y) mediante las fórmulas siguientes:

$$(\rho, \theta) \longrightarrow (x, y) \text{ tal que } x = \rho \cos \theta, \qquad y = \rho \operatorname{sen} \theta;$$

precísense las clases de equivalencia para la relación asociada $\mathcal{F}$. La interpretación geométrica conduce a las *coordenadas polares*.

3. Expresamos una relación de equivalencia $\mathcal{E}$ por un diagrama en coordenadas rectangulares. Demuéstrese que, si el diagrama contiene dos puntos, contiene los vértices del rectángulo obtenido trazando por estos puntos las paralelas a los ejes, cuando uno de los vértices está sobre la diagonal.



4. El diagrama de la figura representa una relación simétrica y transitiva. Compruébese que dicha relación no es una equivalencia. Dedúzcase de ello que es indispensable postular la reflexividad si se quiere una equivalencia.

5. Una relación binaria $\mathcal{C}$ se llama *circular* si $a\mathcal{C}b$ y $b\mathcal{C}c$ implican $c\mathcal{C}a$. Demuéstrese que una relación de equivalencia es circular y reflexiva. Estúdiese el recíproco.

6. Pruébese (cuando proceda) que las siguientes relaciones son de equivalencia y hállense las clases.

a) D//D', para las rectas del espacio.
b) Ox//O'x' y del mismo sentido, para las rectas orientadas.
c) AB equipolente a A'B', para los vectores.
d) El triedro orientado SABC tiene la misma disposición que el triedro orientado S'A'B'C'.
e) El punto M' está alineado con M y con el punto *fijo* O.
f) $\operatorname{tag} \alpha' = \operatorname{tag} \alpha$, para los ángulos.
g) $\operatorname{sen} \alpha = \operatorname{sen} \alpha'$ y $\cos \alpha = \cos \alpha'$.
h) Area ABC = área A'B'C', para los triángulos.
i) La relación $x \equiv y$ (mód. 2), para los enteros relativos.

7. Reconsidérese el último ejercicio del párrafo 10 y hállense, entre las relaciones propuestas, las que son de equivalencia. Determínense las clases.

RELACIONES DE ORDEN

13. Relaciones de orden.—Una relación binaria $\mathcal{R}$ se llama *relación de orden* si es reflexiva, antisimétrica y transitiva.

En otras palabras: para todo x se verifica $x\mathcal{R}x$; $x\mathcal{R}y$ e $y\mathcal{R}x$ implican $x=y$; finalmente $x\mathcal{R}y$ e $y\mathcal{R}z$ implican $x\mathcal{R}z$.

Sobre el conjunto de los números enteros, la relación $x \leqslant y$ es una relación de orden.

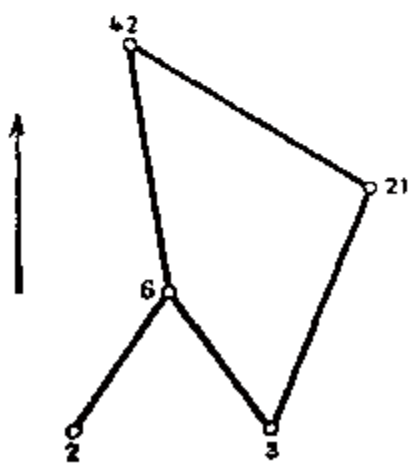
La relación $\leqslant$ caracteriza un orden en sentido amplio, ya que admite la igualdad. La relación $<$ podría llamarse relación de orden estricto, pero nos atendremos a la costumbre actual que tiende a reservar el nombre de relación de orden para las relaciones reflexivas.

Sobre el conjunto de los enteros iguales o superiores a 1, la relación $x \mid y$ (x divide a y) es también una relación de orden.

Se observará, sin embargo, la diferencia siguiente: dos elementos a y b son siempre *comparables* mediante la relación $\leqslant$, ya que siempre tendremos $a \leqslant b$ o bien $b \leqslant a$, e incluso las dos si $a=b$; mientras que dos elementos como 7 y 36 *no son comparables* mediante la relación $\mid$. El hecho de que 7 no divida a 36 no implica que 36 divida a 7.

La relación $\leqslant$ es una relación de *orden total*; la relación $\mid$ es una relación de *orden parcial*.

Diagrama de una relación de orden.—Conviene muchas veces representar por un diagrama un conjunto ordenado mediante una relación de orden total o parcial. Los elementos se simbolizan con puntos o con circulitos (◦); dos elementos comparables se unen con una línea, y, si $a\mathcal{R}b$, colocamos b por encima de a.



La figura adjunta representa el conjunto {2, 3, 6, 21, 42} ordenado por $\mid$.

La relación «a es un ascendiente de b» es una de esas relaciones que pueden llamarse de orden estricto: su diagrama está compuesto de árboles genealógicos. El orden es parcial.

La relación $\subseteq$, definida en el conjunto de las partes de un referencial, es una relación de orden parcial, quizá la más primitiva de todas.

EJERCICIOS:

1. Ordénese por inclusión el conjunto de las partes de $E=\{a, b, c\}$, incluidos ϕ y E. Trácese un diagrama.

2. Para el conjunto de definiciones y teoremas de un libro de texto ponemos $x\mathfrak{P}y$ si x interviene en la demostración o en la definición de y (según el método seguido en el texto). Constrúyase el diagrama de nuestra teoría de relaciones.

3. Sobre la hipotenusa y sobre los catetos de un triángulo rectángulo como diámetros se construyen tres semicircunferencias (lúnulas de Hipócrates). Ordénense por inclusión los diferentes dominios. Diagrama.

4. *Conjuntos casi-ordenados.*—Un conjunto E se llamará casi-ordenado por la relación $\mathfrak{Q}$ si esta relación $\mathfrak{Q}$ es reflexiva y transitiva; el hecho de que $x\mathfrak{Q}y$ e $y\mathfrak{Q}x$ se verifiquen simultáneamente no implica, necesariamente, que $x=y$.

Ejemplo: en el conjunto de los puntos de la superficie terrestre la relación «la altitud de x es menor o igual que la de y».

En estas condiciones, escribimos $x\mathfrak{C}y$ si, y solo si, se verifican a la vez $x\mathfrak{Q}y$ e $y\mathfrak{Q}z$. Pruébese que $\mathfrak{C}$ es una relación de equivalencia. ¿Cuáles son las clases para el ejemplo citado?

5. En el conjunto de los infinitésimos referidos a un infinitésimo principal x, la relación $z \ll y$ (z es infinitésimo con relación a y) no es una relación de orden total porque $y=x$ y $z=x \operatorname{sen} 1/x$ no son comparables.

Determínese en este conjunto una relación de casi-orden $\mathfrak{Q}$ tal que las clases de la relación asociada $\mathfrak{C}$ comprendan en particular los *órdenes de infinitud*.

14. Mayorante. Límite superior.—Sea E un conjunto dotado de una relación de orden. Si el orden es total emplearemos el signo $\leqslant$, que leeremos *inferior o igual a*; si el orden es parcial, utilizaremos el signo $\mathfrak{P}$, que leeremos *precede* o *es anterior a*.

Si $a \leqslant b$, o si $a\mathfrak{P}b$, diremos también que b es un *mayorante* de a.

Si A es una parte de E tal que para todo $a \in A$ se verifica $a \leqslant b$, o $a\mathfrak{P}b$, decimos que b es *un mayorante de* A.

En el caso de orden total, si el conjunto de todos los mayorantes de A contiene un elemento β menor que todos los restantes, se demuestra fácilmente que β es único y lo llamamos *el* límite superior de A.

Las nociones de *minorante* y de *límite inferior* se deducen de las anteriores invirtiendo el orden.

En el conjunto de los enteros, todo subconjunto finito admite un límite superior y otro inferior que pertenecen al subconjunto.

En el conjunto de los enteros naturales, todo subconjunto admite un límite inferior y lo contiene.

En el conjunto de los racionales, el subconjunto A de elementos de la forma $1-1/n$ con n entero y positivo admite el 1 como límite superior, pero no lo contiene: es *abierto*. Añadiéndole el 1 obtenemos un conjunto *cerrado* $\bar{A}$.

El conjunto de los números *racionales* x tales que $x^2 < 2$ admite como mayorantes 1,5 ó 1,42 ó 1,415...; pero se demuestra en aritmética que tal conjunto no admite límite superior (en el conjunto de los racionales).

EJERCICIOS:

1. Demuéstrese que, para el orden total, si el subconjunto A admite un límite superior β, este es único (reducción al absurdo).

2. ¿Por qué no hemos definido el límite superior en el caso de un orden parcial?

3. Consideramos el número decimal de infinitas cifras 0,1234567891011 12..., cuya ley de formación es evidente. Pruébese que el conjunto de *racionales* 0,1; 0,12; 0,123; etc., carece de límite superior (en el conjunto de los racionales).

EJERCICIOS SOBRE EL CAPITULO I

1-1. Cardinales.—Sea $\mathcal{F}$ un conjunto de conjuntos (diremos mejor una *familia* de conjuntos): $\mathcal{F}=\{E, E', E'', \ldots\}$, finita o infinita. Escribiremos $E \sim E'$ si, y solo si, existe una biyección de E en E'.

a) Demuéstrese que $\sim$ es una relación de equivalencia sobre $\mathcal{F}$. Cada clase se llamará un *número cardinal* o un *cardinal* y diremos que dos conjuntos de una misma clase tienen la misma *potencia*.

b) Compruébese que el conjunto de los días de la semana y el conjunto de las maravillas del mundo antiguo tienen la misma potencia; ¿cuál es su cardinal?

c) Cuando un conjunto es aplicable biunívocamente sobre el conjunto N^* de los enteros naturales, se dice que es *numerable* o que su cardinal es $\aleph_0$ (alef de índice cero). Demuéstrese que el conjunto de los números pares, el conjunto de los cuadrados perfectos y el conjunto de los números primos son conjuntos numerables. Dedúzcase de lo anterior que un conjunto infinito puede tener la misma potencia que una de sus partes propias.

d) Fórmese un cuadro doblemente infinito que representa a $N^* \times N^*$; sea (a, b) un elemento. Disponiendo los elementos de este cuadro según las diagonales que corresponden a $a+b=$cte. (2, 3, ...), pruébese que podemos numerar sus elementos. Dedúzcase como consecuencia que el conjunto de los números racionales es numerable.

e) Demuéstrese que el conjunto de las ecuaciones $ax^2+bx+c=0$ de coeficientes enteros relativos es numerable. (Recuérdese el método precedente.) Para empezar se podrá introducir la altura de la ecuación, $H=|a|+|b|+|c|$, y clasificarlas por alturas.

1-2. Ordinales. Sea $\mathcal{O}$ una familia de conjuntos totalmente ordenados $\mathcal{O}=\{T, T', T'', \ldots\}$. Escribiremos $T \sim T'$ si, y solo si, existe una aplicación

biunívoca de T sobre T' que conserve el orden; es decir, que si $a \leqslant b$, se deberá tener $a' \leqslant b'$ para las imágenes.

a) Demuéstrese que $\sim$ es una relación de equivalencia. Llamaremos a cada clase un *tipo de orden*.

b) Compruébese que el conjunto de los enteros y el conjunto de los racionales ordenados según el orden natural no son del mismo tipo.

c) El conjunto ordenado formado por el conjunto de todos los enteros y de un camello colocado detrás de todos ellos no tiene el mismo tipo que el conjunto formado por un camello seguido, a continuación, por los enteros. ¿Cuál es en consecuencia el tipo de este último conjunto?

d) Un conjunto se dice *bien ordenado* si todo subconjunto no vacío admite un primer elemento. N, ordenado naturalmente, está bien ordenado. ¿Está bien ordenado el conjunto de los enteros negativos? La misma cuestión para el conjunto de los x racionales tales que $0 < x < 1$; para el conjunto de las fracciones $1/n$ siendo $n \in \mathbf{N}^*$; para el conjunto de los números $1 - 1/n$.

e) La clase de un conjunto bien ordenado es un *tipo de buen orden* o un *ordinal*. La clase de $\mathbf{N}^*$ es el primer *ordinal transfinito*, que designamos por ω. Defínanse

$$\omega+1,\ \omega+2,\ \ldots,\ \omega+n;\ \omega+\omega,\ \ldots;\omega+\omega+\omega+\ldots+\omega+\ldots=\omega\times\omega.$$

1-3. Doscientos húsares (todos de tallas diferentes) están formados en 20 columnas y 10 filas. Tomamos nota del más alto de cada columna y separamos el menor de los 20; lo designamos por X. Nos fijamos en el más pequeño de cada fila y separamos el mayor de los 10; lo denominamos Y. De X e Y, ¿cuál es el más alto?

CAPITULO II

LEYES DE COMPOSICION, ESTRUCTURAS, ALGEBRA DE BOOLE

LEYES DE COMPOSICION

1. Definición.—Una ley de composición interna definida sobre un conjunto E es una operación que permite hacer corresponder a ciertos pares ordenados (a, b) de elementos, distintos o no, de E un elemento bien determinado del mismo conjunto E. El elemento obtenido de esta forma se llama compuesto de a por b para la ley en cuestión.

Si la ley puede aplicarse a todos los pares (a, b), se dice que está *definida en todo el conjunto*. Una ley de composición aparece, pues, como una función definida sobre una parte de $E \times E$ y que toma sus valores en E.

La adición es una ley de composición definida sobre todo el conjunto de los enteros naturales; el compuesto, que se escribe $a+b$, se llama suma de a y de b. La sustracción es una ley definida sobre el conjunto de pares (a, b) tales que $a \geq b$; el compuesto, denotado por $a-b$, se denomina diferencia. La multiplicación da el producto $a \cdot b$ o ab. La elevación a potencias hace corresponder al par (a, b) el elemento a^b.

En geometría elemental se define el producto de dos transformaciones puntuales efectuadas en un orden determinado. No insistimos sobre este ejemplo, que volveremos a considerar y generalizar a propósito del producto de dos aplicaciones.

Las leyes citadas anteriormente se simbolizan con diversos signos, y en ciertos casos mediante la ausencia de signo (producto) o por una disposición tipográfica particular (potenciación). Cuando tengamos que razonar sobre una ley cualquiera la designaremos por medio del signo $*$. El compuesto de a por b se escribirá $a * b$, que se lee: «a estrella b».

Ley externa.—Una ley de composición externa, definida sobre un conjunto $E = \{a, b, c, \ldots\}$ por medio de un *dominio de operadores* $\Omega = \{\alpha, \beta, \gamma, \ldots\}$, es una operación que permite hacer corresponder a todo par ordenado de $\Omega \times E$ un elemento bien determinado de E.

Tal ley aparece, pues, como una función que toma sus argumentos en E×E y sus valores en E. El ejemplo más sencillo lo proporciona el producto de vectores, libres o fijos, cuando se toman los números reales como operadores.

Parte cerrada o estable para una ley.—Una parte A de E es *cerrada* (o *estable*) para la ley $*$ si el compuesto de dos elementos de A pertenece siempre a A. De esta forma, A aparece dotado de una ley de composición que llamamos restricción de $*$ a A.

$*$	a	b	c	d
a				
b				
c				
d				

Caso de un conjunto finito: tabla de una ley.—Si E es finito, es posible *tabular* una ley $*$. Construimos una tabla cuadrada de doble entrada y consignamos el valor del elemento $a*b$ en la casilla situada en la intersección de la fila de a (factor de la izquierda) con la columna de b (factor de la derecha). Escribimos el signo de la ley tabulada en el ángulo superior izquierdo. La tabla de Pitágoras es el venerable ascendiente de estas *tablas de composición*.

EJERCICIOS:

1. $\{0\}$, $\{0, 1\}$, $\{1\}$, $\{-1, 1\}$, $\{-1, 0, 1\}$ son partes del conjunto de los enteros cerradas para la multiplicación.

2. Examínese el conjunto de los números pares para las leyes $a+b$, $a\times b$, a^b.

3. En el conjunto de los enteros relativos designamos por $a*b$ el menor de los números a o b. ¿Cuáles son las partes cerradas?

4. Sobre el conjunto $E=\{1, 2, 3, 6\}$ designamos por $a-b$ al m.c.d. de a y b. Tabúlese esta ley.

5. Sobre el conjunto $E=\{1, 2, 5\}$ se designa con $a*b$ el resto de la división de a^b por 3. Tabúlese esta ley. Explíquese por qué obtenemos una ley de composición interna.

6. ¿Es interna la ley E×F definida sobre la familia de conjuntos {E, F, G}?

2. Extensión de una ley de E a E×E.—Definida una ley $*$ sobre E, es posible dotar a los pares ordenados (a, b) que forman E×E de una ley simbolizada por el mismo signo, y tal que

$$(a_1, b_1) * (a_2, b_2) = (a_1 * a_2,\ b_1 * b_2).$$

Tenemos un ejemplo en la adición de vectores libres (o de vectores fijos de origen O) en un plano euclídiano referido a un sistema de ejes. Un

vector está representado por el par ordenado de sus componentes escalares, y la adición definida sobre el conjunto R de los números se extiende al conjunto producto $R \times R$, conjunto de los vectores.

La ley definida de esta manera se llama, por ciertos autores, *ley producto,* designación que desgraciadamente se presta a equívoco.

$*$	0	1
0	0	1
1	1	0

EJERCICIOS:

1. La multiplicación de fracciones puede definirse a partir de una ley producto.

2. Sea $E=\{0, 1\}$ dotado de la ley tabulada en el margen izquierdo de la página.

Escríbase la tabla de la ley $*$ extendida a $E \times E$.

3. Ley estable respecto de una equivalencia.—Consideremos un conjunto E dotado simultáneamente de una relación de equivalencia $\mathcal{E}$ y de una ley de composición $*$.

Estabilidad.—Decimos que la ley $*$ es *estable* para la equivalencia $\mathcal{E}$ si el compuesto de dos elementos cualesquiera a y b se cambia en un equivalente cuando se reemplazan a y b por sus equivalentes. En otras palabras:

$$a \mathcal{E} a' \quad \text{y} \quad b \mathcal{E} b' \Longrightarrow (a * b)\ \mathcal{E}(a' * b').$$

TEOREMA.—*Una ley de composición* $*$, *definida sobre un conjunto* E *y estable respecto de una relación de equivalencia* $\mathcal{E}$, *induce sobre el conjunto cociente* $E/\mathcal{E}$ *una ley de composición* $\bar{*}$.

Sean $C_1=C(a)$ y $C_2=C(b)$ dos clases de equivalencia cualesquiera. Demostremos que podemos asociarlas una clase bien determinada C_3. Elijamos $a \in C_1$ y $b \in C_2$ y consideremos el elemento $c=a*b$, que pertenece a una clase $C_3=C(c)$. La estabilidad de la ley $*$ prueba que C_3 es independiente de los elementos a y b elegidos en C_1 y C_2; la clase C_3 solo depende, pues, de los C_1 y C_2 elegidos, lo que define una ley de composición sobre el conjunto $E/\mathcal{E}$; si la simbolizamos por $\bar{*}$, tendremos:

$$C_1 \bar{*} C_2 = C_3.$$

Indiquemos una aplicación bien conocida del lector: la prueba del 9. Sobre el conjunto N^* de los enteros definimos una relación de equivalencia: un par (a, b) la verifica si, y solo si, la diferencia de los números que lo constituyen (el mayor—el menor) es un múltiplo de 9. El conjunto de los

números de la forma $a \pm 9k$ constituye una clase, siendo a un elemento cualquiera de ella. Toda clase contiene un elemento y uno solo $0 \leqslant a < 9$, y hay nueve clases. Utilizando el sistema de base diez, obtenemos a observando que cualquier número pertenece a la misma clase que la suma de sus cifras y reiterando esta regla cuantas veces sea necesario.

Además, se verifica fácilmente que la adición y la multiplicación son estables respecto de la equivalencia precedente, lo que induce ciertas operaciones sobre el conjunto de las clases. La prueba del 9 consiste justamente en comprobar la estabilidad de la operación: p. ej., la clase del producto debe ser el producto de las clases de los factores. Esto se hace representando cada clase por el elemento $0 \leqslant a < 9$ que la caracteriza; de esta forma se obtiene la clásica regla.

EJERCICIOS:

1. En el conjunto de los enteros relativos, las leyes + y × son estables respecto de la equivalencia $|a-b| = \text{mult. } 2$.

Demuéstrese que las operaciones inducidas son:

par + par = par,	par × par = par,
par + impar = impar,	par × impar = par,
impar + impar = par,	impar × impar = impar.

2. Pruébese que la adición geométrica de vectores del mismo origen es una operación estable respecto de la equipolencia. ¿Qué consecuencia se deduce para los vectores libres?

PROPIEDADES USUALES DE LAS LEYES DE COMPOSICION

4. Leyes asociativas.—Formado el elemento $a * b$, podemos componerlo con un tercer elemento c; de esta forma se obtiene un resultado que escribimos $(a * b) * c$, utilizando los paréntesis para indicar el resultado de una operación efectuada.

Existe otra manera (y una sola) de agrupar los mismos *factores* tomados en el mismo orden; conduce al resultado $a * (b * c)$, que no es necesariamente igual al precedente.

Para n elementos $a_1, a_2, \ldots, a_n$ tomados en un orden determinado, el compuesto se define de manera natural componiendo primero los dos primeros elementos, el resultado con el tercero, el resultado con el cuarto, etc. El empleo de paréntesis nos permite escribir así este producto:

$$\Big(\ldots \big((a_1 * a_2) * a_3 \big) * \ldots \Big) * a_n.$$

Pero con los mismos elementos, considerados siempre en el mismo

orden, podremos formar agrupaciones diferentes, que se traducen por distintas distribuciones de los paréntesis. Se facilita la lectura si se coloca un subíndice cuando se abre el paréntesis y el mismo subíndice cuando se cierra. Así, p. ej.,

$$\Big(a_1 * \Big(\big(a_2 * a_3\big)_1 * \big(a_4 * a_5\big)_2\Big)_3\Big)_4 * \big(a_6 * a_7\big)_5 ;$$

la estrella de la última operación es la única que queda fuera de todo paréntesis.

Asociatividad.—Se denomina *asociativa* una ley tal que

$$(a * b) * c = a * (b * c)$$

cualesquiera que sean a, b y c. El resultado común se escribe sin paréntesis $a * b * c$.

Es notable que la unicidad se extiende a n factores sin más que la hipótesis anterior; precisemos esta propiedad:

TEOREMA DE ASOCIATIVIDAD.—*Para una ley asociativa el compuesto de n elementos* $a_1, a_2 \ldots, a_n$ *tomados en este orden es único; no depende de la manera en que se agrupen los elementos en el transcurso del cálculo.*

Por convenio, el resultado, $a_1 * a_2 * \ldots * a_n$ se escribe sin paréntesis.

El teorema se cumple por definición para $n=3$. Supongámoslo cierto para todo conjunto ordenado de $n-1$ elementos y consideremos, para n elementos, uno cualquiera de los órdenes de composición posibles.

La estrella que indica la última operación separa un par de paréntesis escrito a la izquierda y un par análogo escrito a la derecha; uno y otro pueden comprender solo un elemento. En cualquier caso, cada par contiene menos de n elementos, y, la hipótesis de recurrencia permite suprimir los paréntesis. Llegamos, pues, bien a

$$a_1 * (a_2 * a_3 * \ldots * a_{n-1} * a_n) \text{ (caso } \alpha),$$

o bien a

$$(a_1 * a_2 * \ldots * a_k) * (a_{k+1} * \ldots * a_{n-1} * a_n) \text{ (caso } \beta),$$

o finalmente a

$$(a_1 * a_2 * \ldots * a_{n-1}) * a_n \text{ (caso } \gamma).$$

El caso γ corresponde a la definición natural del producto. Demostraremos que los casos α y β se reducen al anterior; p. ej., β da sucesivamente:

$$(a_1 * a_2 * \ldots * a_k) * [(a_{k+1} * \ldots * a_{n-1}) * a_n]$$

por la hipótesis de recurrencia;

$$[(a_1 * a_2 * \ldots * a_k) * (a_{k+1} * \ldots * a_{n-1})] * a_n$$

por el axioma de asociatividad, y finalmente

$$(a_1 * a_2 * \ldots * a_{n-1}) * a_n$$

por la hipótesis de recurrencia. Se obtiene así la tercera expresión (caso γ). El caso α se demuestra de manera análoga.

Cualquiera que sea la forma de agrupar los factores, se obtiene el mismo resultado que con la definición natural de producto. El teorema es, pues, cierto para n elementos.

Las leyes que generalmente estudiamos en matemática son asociativas. Cabe, sin embargo, señalar la excepción del producto vectorial

$$(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} \neq \mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}).$$

5. Otras propiedades usuales. Conmutatividad.—Sean a y b dos elementos cualesquiera; en general, los elementos $a * b$ y $b * a$ son distintos.

Se dirá que una ley es *conmutativa* cuando $a * b = b * a$, cualesquiera que sean a y b.

La conmutatividad considerada en sí misma carece casi de interés. Por el contrario, combinada con la ley asociativa da un resultado sencillo y útil.

TEOREMA.—*Para una ley asociativa y conmutativa, el compuesto de n elementos es independiente del orden de sus términos.*

Sea

$$P = a_1 * a_2 * \ldots * a_{i-1} * a_i * a_{i+1} * \ldots * a_n$$

A) Se pueden permutar dos elementos consecutivos a_i y a_{i+1}. En efecto: hagamos $p = a_1 * a_2 * \ldots * a_{i-1}$ y $q = a_{i+2} * \ldots * a_n$; se tiene:

$$\begin{aligned} P &= p * a_i * a_{i+1} * q && \text{(asociatividad),} \\ &= [p * (a_i * a_{i+1})] * q && \text{(asociatividad),} \\ &= [p * (a_{i+1} * a_i)] * q && \text{(conmutatividad),} \\ &= p * a_{i+1} * a_i * q && \text{(asociatividad),} \end{aligned}$$

por consiguiente,

$$= a_1 * a_2 * \ldots * a_{i-1} * a_{i+1} * a_i * a_{i+2} * \ldots * a_n$$

por asociatividad.

B) Se pueden colocar los a_k en un orden asignado de antemano. En efecto: llevaremos a_1 a su lugar por una sucesión de permutaciones de dos elementos consecutivos; después haremos lo mismo con a_2, etc.

Distributividad.—Supongamos un conjunto dotado de dos leyes de composición denotadas, respectivamente, por $\circ$ y $*$. Diremos que la ley $\circ$ es *distributiva a la izquierda* con relación a la ley $*$ si se verifica siempre que

$$a \circ (b * c) = (a \circ b) * (a \circ c).$$

Se define igualmente la distributividad a *la derecha,* denominándose sencillamente *distributiva* la distributividad bilateral. Si $\circ$ es distributiva y $*$ asociativa, tenemos (como es fácil probar):

$$(a_1 * a_2) \circ (b_1 * b_2) = (a_1 \circ b_1) * (a_2 \circ b_1) * (a_1 \circ b_2) * (a_2 \circ b_2).$$

En el caso de los enteros positivos tenemos, p. ej.,

$$(a_1 + a_2) \cdot (b_1 + b_2) = a_1 \cdot b_1 + a_1 \cdot b_2 + a_2 \cdot b_1 + a_2 \cdot b_2,$$

donde por costumbre se han suprimido los paréntesis en el segundo miembro.

Sobre el conjunto de vectores se tiene también la fórmula

$$(\alpha_1 + \alpha_2)(\mathbf{V}_1 + \mathbf{V}_2) = \alpha_1 \mathbf{V}_1 + \alpha_1 \mathbf{V}_2 + \alpha_2 \mathbf{V}_1 + \alpha_2 \mathbf{V}_2$$

que resulta de una doble distributividad de la ley externa con relación a la adición de vectores y a la adición de escalares.

Caso de una ley inducida.—Precisemos las propiedades de la ley inducida definida en la sección 3.

TEOREMA.—*Sea, sobre un conjunto, una ley de composición estable respecto de una equivalencia. Una ley asociativa induce sobre el conjunto cociente una ley asociativa. Una ley conmutativa induce una ley conmutativa. Dos leyes estables ligadas por distributividad inducen dos leyes ligadas por la misma distributividad.*

Estas propiedades resultan fácilmente de la observación siguiente: toda relación de igualdad entre diversos elementos, compuestos según la ley de E, se conserva cuando pasamos a la ley inducida sobre el conjunto cociente.

EJERCICIOS:

1. ¿Está la ley $a * b = \frac{1}{2}(a+b)$ definida sobre todo el conjunto de los enteros naturales? ¿Es conmutativa? ¿Es asociativa?

2. En el conjunto de las homotecias del mismo centro O, se define la *suma* de dos homotecias como otra homotecia del mismo centro y razón de homotecia la suma de las razones. Estúdiese cómo se comporta el producto de homotecias con relación a esta adición.

3. La tabla adjunta define la ley $*$ sobre

$$E = \{a, b, c\}.$$

$*$	a	b	c
a	a	c	b
b	c	b	a
c	b	a	c

Demuéstrese que es conmutativa pero no asociativa.

Comparando $(a * b) * c$ y $(a * c) * b$ es fácil convencerse de la poca eficacia de la conmutatividad cuando actúa ella sola.

4. Sobre el conjunto de los enteros mayores o iguales a 1 consideramos dos leyes de composición:

$a \frown b$ designa el m.c.d. de a y b
$a \frown b$ designa el m.c.m. de a y b.

Estúdiense desde el punto de vista de la conmutatividad y de la asociatividad.

Pruébese que la multiplicación es distributiva con relación a estas dos leyes.

5. Sobre el conjunto de los racionales positivos ponemos $a * b = a + 1/b$. Demuéstrese que esta ley no es asociativa y fórmense todos los compuestos obtenidos con a, b, c y d en este orden.

6. Precísese con detalle la demostración del último teorema de este parágrafo.

7. Pruébese que si una ley estable de E admite un elemento neutro e (véase más adelante), la ley inducida admite uno que es la clase $C(e)$.

ELEMENTO NEUTRO, ELEMENTO REGULAR

6. Algunas propiedades posibles de los elementos.—Las siguientes definiciones son de constante empleo.

Elemento neutro.—Decimos que un elemento e es *neutro* para la ley $*$ si, para todo a,

$$a * e = e * a = a.$$

El número 0 es neutro para la adición de los enteros, el número 1 es neutro para la multiplicación.

TEOREMA.—*Si existe, el elemento neutro es único.*

Si fuesen e y e', los dos, elementos neutros, se tendría $e * e' = e'$, ya que e es neutro; pero $e * e' = e$ por ser e' neutro, luego $e = e'$.

Elemento regular.— La igualdad $b = c$, que expresa la identidad de b y de c, implica evidentemente $a * b = a * c$ y $b * a = c * a$, cualquiera que sea a.

Pero no es posible decir que el recíproco sea cierto. Por eso, cuando $a * b = a * c$ permite afirmar, en todos los casos, que $b = c$, se dice que el elemento a es *regular a la izquierda*. De la misma manera se define la *regularidad a la derecha*. Llamamos elemento *regular* al que lo es a la derecha y a la izquierda simultáneamente. Cuando todos los elementos son regulares se dice que *es válida la regla de simplificación* sobre el conjunto para la ley considerada.

Para la adición de enteros todos los elementos son regulares. Para la multiplicación, existe la excepción del número cero, ya que de $0 \cdot a = 0 \cdot b$ no es posible deducir nada.

7. Elementos simétricos.—Sea e el elemento neutro de una ley $*$. Decimos que dos elementos a y a' son simétricos cuando

$$a * a' = a' * a = e.$$

En el conjunto de los enteros relativos el simétrico de a con relación a la ley aditiva es su *opuesto* $(-a)$. En el conjunto de los racionales no nulos el simétrico de a es su *inverso* $1/a$.

De manera general, el simétrico de a, si existe, se designa por a^{-1}, y la reciprocidad entre a y a' que implica la propia definición prueba que $(a^{-1})^{-1} = a$.

TEOREMA.—*Para una ley asociativa, si el elemento a admite un simétrico, este es único.*

Sean a^{-1} y a' dos simétricos del elemento a. Se tendrá:

$$a * a^{-1} = a^{-1} * a = e \qquad \text{y} \qquad a * a' = a' * a = e.$$

Componiendo a la izquierda los dos miembros de $a * a' = e$ con a^{-1}, resulta:

$$a^{-1} * (a * a') = a^{-1} * e = a^{-1},$$
$$(a^{-1} * a) * a' = a^{-1},$$

de donde

$$e * a' = a^{-1}, \qquad a' = a^{-1}.$$

Por consiguiente, no existe más que un simétrico.

TEOREMA.—*Para una ley asociativa, si a y b admiten simétricos el compuesto $a * b$ también lo admite y viene dado por la fórmula*

$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

Tenemos, en efecto,

$$\begin{aligned}(a * b) * (b^{-1} * a^{-1}) &= a * (b * b^{-1}) * a^{-1} \\ &= a * e * a^{-1} = a * a^{-1} = e;\end{aligned}$$

análogamente,

$$(b^{-1} * a^{-1}) * (a * b) = e.$$

8. La notación exponencial.—Sea una ley asociativa definida sobre un conjunto E. Si $a \in E$, se escribe

$$a * a = a^2, \qquad a^2 * a = a * a^2 = a^3$$

y, en general,

$$a^{p-1} * a = a^p.$$

A causa de la asociatividad son fáciles de comprobar las clásicas fórmulas

$$a^p * a^q = a^{p+q}, \qquad (a^p)^q = a^{pq} \qquad [1]$$

ya que en ambos miembros encontramos el mismo número de veces el factor a. Indiquemos las generalizaciones clásicas de estas fórmulas, válidas hasta aquí para p y q enteros mayores o iguales a 2.

1.° Conviniendo que $a = a^1$ las extendemos inmediatamente al caso en que p y q toman el valor 1.

2.° Si la ley considerada posee un elemento neutro e, ponemos, para cualquier a, $a^0 = e$. De donde,

$$\begin{gathered}a^p * a^0 = a^0 * a^p = a^{p+0} = a^p, \\ (a^p)^0 = (a^0)^p = a^{0 \cdot p} = e.\end{gathered}$$

3.° Por último, si a posee un simétrico a^{-1}, se pone $(a^{-1})^p = a^{-p}$: el último teorema del apartado anterior nos dice que a^{-p} es el inverso de a^p. Se comprueban finalmente las fórmulas [1] para p y q enteros positivos, negativos o nulos. Si $p' < 0$, comprobamos, p. ej., que

$a^{p'} * a^q = a^{p'+q}$. Haciendo $p' = -p$, esto se escribe $a^{-p} * a^q = a^{q-p}$, de donde

$$(a^{-1} * a^{-1} * \ldots * a^{-1}) * (a * a * \ldots * a) = a^{q-p}.$$

El primer factor consta de p veces a^{-1}, el segundo de q veces a. Aplicando la asociatividad agrupamos los factores a^{-1} y a que estén más próximos, lo que da e. Si $p < q$, quedarán $q - p$ factores a; es decir,

$$a^{-p} * a^q = a^{q-p}.$$

Si $p > q$, quedan $p - q$ factores a^{-1}, de donde

$$a^{-p} * a^q = (a^{-1})^{p-q} = a^{q-p}.$$

La ley se verifica, pues, para $p < 0$.

Dejamos al cuidado del lector el examen de los restantes casos y la demostración de la segunda ley [1].

4.° Si, además, la ley considerada es conmutativa, se tendrá:

$$(a * b)^p = a^p * b^p.$$

EJERCICIOS:

1. Sobre el conjunto de los enteros positivos o nulos adoptamos la ley

$$a * b = a + b + a \cdot b.$$

Examínense la conmutatividad y la asociatividad. ¿Tiene elemento neutro?

2. Sobre un conjunto E, dotado de una ley $*$, e se llama *neutro a la izquierda* si $e * a = a$ para todo $a \in E$; de igual forma, f *es neutro a la derecha* si $a * f = a$. Demuéstrese que si e y f existen simultáneamente, $e = f$.

3. Sobre todo conjunto E la ley definida por $a * b = b$ es claramente una ley interna. Estúdiense la asociatividad y la conmutatividad. Demuéstrese que todo elemento es neutro a la izquierda. Pruébese a partir del ejercicio anterior que no puede existir elemento neutro a la derecha.

4. Sea E un conjunto dotado de una ley asociativa y que posee elemento neutro e. Dado un elemento a, si existe a' tal que $a * a' = e$, se dice que a' es un inverso a la derecha de a. Demuéstrese que, si a es regular, el elemento a' es único y que, si a' es regular, a' es también el único inverso a la izquierda.

Este ejercicio desmenuza la noción global de elementos simétricos.

COMPOSICION DE APLICACIONES

9. Producto de dos aplicaciones.—Sean tres conjuntos cualesquiera E, F, G, y consideremos una aplicación f de E en F y una aplicación g de F en G. Si $a \in E$, f da de a una imagen $b=f(a)$, y g da de b una imagen $c=g(b)$. El paso directo de a a c define una aplicación h de E en G tal que

$$c=h(a)=g[f(a)].$$

Por definición, h se llama *compuesta* o *producto* de las aplicaciones f y g. Simbolizando esta ley de composición por el signo $\circ$, escribimos:

$$h=g \circ f$$

Para respetar la notación funcional nos vemos obligados a escribir las sucesivas aplicaciones de derecha a izquierda. Es interesante señalar que el producto definido de esta manera no es una ley de composición en el sentido dado a este término (véase pág. 22). En efecto, f, g y h no son elementos de un mismo conjunto.

Asociatividad.—La composición definida anteriormente es asociativa. Sean cuatro conjuntos E, F, G y H, f una aplicación de E en F, g una aplicación de F en G, y h una aplicación de G en H. La asociatividad se expresa por la fórmula

$$h \circ (g \circ f)=(h \circ g) \circ f,$$

que resulta de la definición del producto. En efecto, para $a \in E$ el primer miembro da

$$[h \circ (g \circ f)]a=h[(g \circ f)a]=h(g[f(a)])$$

y el segundo

$$[(h \circ g) \circ f]a=(h \circ g)\, f\,(a)=h(g[f(a)]).$$

Es manifiesta la identidad de resultados, cualquiera que sea el elemento de E considerado.

10. Aplicación de un conjunto en sí mismo.—Consideremos el caso particular de que los conjuntos F y G se confunden con E. Las aplicaciones f y g pertenecen al mismo conjunto: el de las aplicaciones de E en E. Este conjunto está bien definido, caracterizándose en él

la igualdad por $f=f'$ si $f(a)=f'(a)$ para todo $a \in E$. La definición anterior asocia a las aplicaciones f y g tomadas en este orden una aplicación producto, $g \circ f$, que pertenece igualmente al conjunto considerado. Tenemos, pues, una ley de composición sobre este conjunto y, según la propiedad precedente, esta ley es asociativa. En general, no es conmutativa, $g \circ f \neq f \circ g$, por lo que deberemos respetar el orden de los factores.

EJEMPLOS.- -El estudio de las transformaciones geométricas proporciona numerosos ejemplos de composición. E es un conjunto de puntos; por lo general, todo el espacio o uno de sus planos. El lector comprobará que los productos de transformaciones clásicas, como simetría, traslación, rotación, homotecia, satisfacen a la definición general precedente.

Elemento neutro.— La ley de composición que acabamos de definir admite un elemento neutro: la aplicación idéntica i que a todo $a \in E$ asocia $i(a)=a$. Se comprueba, en efecto, que para cualquier f

$$i \circ f = f \circ i = f.$$

Simétrico. TEOREMA.—*Una aplicación f de E en sí mismo admite una aplicación simétrica (inversa) para la ley $\circ$ si, y solo si, es biyectiva.*

La condición es suficiente porque toda aplicación biyectiva, f, admite una biyección recíproca, ya simbolizada por f^{-1}, y tal que

$$b=f(a) \quad \text{implica} \quad a=f^{-1}(b), \quad (a, b \in E)$$

de donde

$$f^{-1}[f(a)]=a \quad \text{y} \quad f[f^{-1}(b)]=b;$$

es decir,

$$f^{-1} \circ f = f \circ f^{-1} = i.$$

Ella constituye, pues, el elemento simétrico de f en el sentido dado en la sección 7. Recíprocamente, si existe f' tal que

$$f \circ f' = f' \circ f = i,$$

la aplicación f es biyectiva. La relación $f[f'(a)]=a$ muestra que f es suprayectiva, ya que los transformados por f de los elementos de la forma $f'(a)$ recubren por entero E (a arbitrario). Además, la relación $f'[f(a)]=a$ prueba que f es inyectiva, porque si a y $a'(a \neq a')$ tuvieran

la misma imagen $b=f(a)=f(a')$ por f, el elemento b tendría dos imágenes a y a' por f', lo que contradice la definición de aplicación. De ello resulta que f es biyectiva, siendo f' la biyección recíproca f^{-1}

EJERCICIOS:

1. En geometría plana, *la simetría con relación a una recta* es una aplicación. La designaremos por la misma letra que su eje.
Estúdiense $\Delta \circ \Delta$ y $\Delta_1 \circ \Delta_2$ distinguiendo entre $\Delta_1 // \Delta_2$ y $\Delta_1 \neq \Delta_2$.
¿Qué sucede si cambiamos el orden de los factores?

2. Aplicamos $E=\{a, b, c\}$ en sí mismo, mediante $\begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix}$; hecho esto, volvemos a aplicar E en sí mismo por $\begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}$. ¿Cuál es el producto de estas dos aplicaciones?

3. Sean f y g dos aplicaciones permutables ($f \circ g = g \circ f$) de un conjunto E en sí mismo. Demuéstrese que todo elemento invariante en la transformación f tiene por imagen (mediante g) un elemento invariante mediante f.

NOCIONES DE ESTRUCTURA Y DE ISOMORFISMO

11. Estructura.—Dotar a un conjunto de una o varias relaciones y leyes de composición es conferirle una *estructura*. Una estructura, por consiguiente, queda definida por los axiomas que rigen las relaciones y las operaciones de que está dotada. En lo que sigue estudiaremos las estructuras fundamentales del álgebra: grupos, anillos, cuerpos y espacios vectoriales.

Los axiomas de una estructura no surgen del capricho del matemático, sino que son consecuencias necesarias del desarrollo de la ciencia y adquieren permanencia por su generalidad y fecundidad.

El estudio de una estructura puede hacerse de manera «abstracta», dejando indeterminada la naturaleza de sus elementos.

Podemos también dar realidad a esta estructura por medio de elementos que psicológicamente aparecen como *concretos* (números, transformaciones geométricas), ya que las abstracciones de donde provienen son anteriores en la historia de la ciencia.

Estos dos puntos de vista se complementan en una sana pedagogía y el lector habrá comprobado que los adoptamos sucesivamente: el texto en caracteres gruesos cuida el rigor lógico cuanto le es posible, mientras que los comentarios y ejemplos ilustran y justifican intuitivamente el origen de las nociones abstractas.

Los enteros naturales.—No juzgamos útil repetir aquí la axiomática de los enteros naturales; el lector interesado puede consultar otros textos de matemáticas. Pero las restantes estructuras las presentaremos *ab initio*.

EJERCICIO

Formúlese la estructura del conjunto N^* de los enteros naturales utilizando el lenguaje de la teoría de las relaciones y de la teoría de las leyes de composición. No se pretenda reducir los axiomas al mínimo, pero no se olviden las propiedades de «enlace» tales como

$$a < b \qquad \text{implica} \qquad a + x < b + x.$$

12. Isomorfismo.—Sean dos conjuntos equipotentes E y E', y f, una biyección de E sobre E'. Supongamos definidas sobre E una ley de composición denotada por $*$ y sobre E' otra ley simbolizada por $*'$. La biyección f se llama *regular* respecto de las leyes $*$ y $*'$ si la igualdad $a * b = c$, supuesta verificada en E, implica para las imágenes $f(a)$, $f(b)$ y $f(c)$ la igualdad

$$f(a) *' f(b) = f(c)$$

que se verifica sobre E'. Esto equivale a decir que el transformado del compuesto sobre E de dos elementos es el compuesto sobre E' de los transformados de estos elementos.

De ello resulta que a toda igualdad entre diversos elementos de E compuestos por la ley $*$ corresponde una igualdad entre sus imágenes en E', compuestos de la misma manera según la ley $*'$. Por consiguiente, toda propiedad de la ley $*$ (asociatividad, conmutatividad, elemento neutro, inverso) pertenece igualmente a la ley $*'$ de E'. Los dos conjuntos poseen, pues, propiedades algebraicas idénticas. Son *isomorfos* y pueden considerarse como dos realizaciones concretas de la misma estructura abstracta. La noción de isomorfismo es uno de los conceptos básicos del álgebra moderna.

La noción se extiende además a estructuras más ricas. Sobre E y E' puede existir un mismo número de leyes de composición; a toda ley de E debe corresponder una ley de E' tal que f sea regular para cada par. Cabe también que existan relaciones binarias: $\mathcal{R}$ sobre E, $\mathcal{R}'$ sobre E'; la biyección f se llama regular respecto de $\mathcal{R}$ y $\mathcal{R}'$ si

$$a \mathcal{R} b \qquad \text{implica} \qquad f(a) \mathcal{R}' f(b).$$

Las relaciones $\mathcal{R}$ y $\mathcal{R}'$ gozan entonces de las mismas propiedades. Los

dos conjuntos E y E' se llaman isomorfos para las estructuras definidas por las leyes y relaciones consideradas. Indiquemos algunos ejemplos.

Primer ejemplo.—Los símbolos 0, 1, 10, 11, 100, ..., empleados en numeración binaria y los símbolos 0, 1, 2, 3, 4, 5, ..., del sistema decimal forman dos conjuntos isomorfos para las leyes $+$, $\cdot$, la relación $<$, etc. Estos dos conjuntos «concretos» son isomorfos con el conjunto «abstracto» de los enteros, del que dan dos representaciones.

Segundo ejemplo.—En matemáticas elementales nos familiarizamos, en un espíritu esencialmente práctico, con el isomorfismo que liga el conjunto de los números estrictamente positivos al conjunto de todos los números relativos, isomorfismo cuyo «léxico» es la tabla de logaritmos y cuya «gramática» tiene las reglas

$\times$ corresponde a $+$
$:$ corresponde a $-$
elevar a la potencia n corresponde a multiplicar por n
extraer la raíz n-ésima corresponde a dividir por n
$\leqslant$ corresponde a $\leqslant$

Volveremos más adelante sobre este isomorfismo, que es de importancia capital. Pero antes de ello nos encontraremos otros.

Tercer ejemplo.—Para definir los números negativos se introducen a veces dos clases de números afectando de un signo calificativo, $+$ o $-$, a los números hasta entonces en uso. Termina la teoría identificando los números positivos con los números ordinarios. Para evitar que este procedimiento parezca un juego de prestidigitación, conviene señalar: *a)* que un número afectado de un signo no es ya un número ordinario; *b)* que el conjunto de los números precedidos del signo $+$ es *isomorfo* con el de los números ordinarios; *c)* que la asimilación es un *convenio*, cuya posibilidad se basa en dicho isomorfismo.

La construcción de los números racionales da lugar a una observación análoga para los símbolos $n/1$. No insistimos porque esta construcción se estudiará más adelante.

OBSERVACIÓN.—Supongamos que existe una biyección f entre dos conjuntos E y E'. Si E posee ciertas leyes de composición, es posible definir leyes correspondientes sobre E'. A toda ley $*$ sobre E asociamos una ley $*'$ sobre E' definiendo el compuesto $c'=a'*'b'$ de los elementos $a'=f(a)$ y $b'=f(b)$ por $c'=f(c)$, siendo $c=a*b$. Los dos

conjuntos resultan de este modo isomorfos. Decimos que las operaciones sobre E' se han definido por isomorfismo.

Automorfismo.—Sea un conjunto E sobre el que está definida una ley $*$. Si es f una biyección de E en sí mismo, regular para esta ley, nos hallamos ante un caso particular de isomorfismo, debido a que los dos conjuntos E y E' están confundidos. El isomorfismo toma entonces el nombre de *automorfismo*.

En geometría euclidiana, p. ej., la simetría con relación a un plano es un automorfismo del conjunto de los vectores libres del espacio dotado de la ley clásica de adición.

EJERCICIOS:

1. Demuéstrese que el conjunto de los lados de un triángulo es isomorfo con el conjunto de sus ángulos para la relación de orden $\leqslant$.

2. Pruébese que el conjunto de las operaciones geométricas que transforman un rectángulo en sí mismo es isomorfo con el conjunto análogo definido a propósito de un rombo.

La misma cuestión asociando a un paralelepípedo rectángulo otro poliedro.

3. Sea una familia de conjuntos {E, E', E'', ...}. Demuéstrese que la relación «... es isomorfo a ...» es una relación de equivalencia.

4. Los axiomas de Peano (consúltese un libro de aritmética) no definen *los enteros*, sino cualquier conjunto isomorfo con el de los enteros.

5. El conjunto de los vectores libres dotados de la adición es isomorfo con el conjunto de las traslaciones dotado de la ley $\circ$.

6. Determínese el valor (teórico) que toma la fracción $\dfrac{x^3-x+1}{x^2+1}$ cuando se reemplaza x por $2+\sqrt{3}$. Seguidamente, sin nuevos cálculos, hállese su valor para $x=2-\sqrt{3}$. Explíquese empleando la noción de automorfismo.

13. Imagen homomorfa de un conjunto dotado de una ley.—Consideremos dos conjuntos E y E' dotados respectivamente de leyes de composición $*$ y $*'$; supongamos que existe una aplicación suprayectiva f de E sobre E' tal que la imagen del compuesto de a por b mediante $*$ se superpone sobre el compuesto de a' por b' mediante $*'$; de manera más precisa:

$$a \longrightarrow a' \quad \text{y} \quad b \longrightarrow b' \quad \text{y} \quad a * b = c \quad \text{y} \quad a' *' b' = c' \quad \text{implican} \quad c \longrightarrow c'.$$

Decimos entonces que E' es una *imagen homomorfa* de E, dando a f el nombre de *homomorfismo*. Bien entendido, en el caso en que f sea una biyección, este concepto se confunde con el de isomorfismo.

Sabemos que la aplicación f de E sobre E' da origen sobre E a

una equivalencia asociada $\mathcal{F}$ en la que los elementos a, a_1, a_2, ... de una clase $C(a)$ son los que tienen la misma imagen a', de suerte que el conjunto cociente $E/\mathcal{F}$ se aplica biunívocamente sobre el conjunto imagen E' por la regla $C(a) \Longleftrightarrow a'$.

Por otra parte, $\mathcal{F}$ induce sobre $\overline{E} = E/\mathcal{F}$ una ley $\bar{*}$, y decir que $C(a) \bar{*} C(b) = C(c)$ o que $a' *' b' = c'$ equivale evidentemente a traducir el mismo contenido matemático en términos ligeramente diferentes.

De manera rigurosa, el conjunto $\overline{E}$ dotado de la ley $\bar{*}$ es isomorfo con E' dotado de $*'$.

Teorema de homomorfismo.—*Sea* E *un conjunto dotado de una ley de composición* $*$. *Si* E' *es una imagen homomorfa de* E *por la aplicación* f, *la equivalencia* $\mathcal{F}$ *asociada a* f *determina un conjunto cociente* $E/\mathcal{F}$ *que es isomorfo con* E' *dotado de la ley* $*'$.

Añadamos que si E' es una parte de E, el homomorfismo toma el nombre de *endomorfismo.*

Caso en que existe elemento neutro.—Si existe un elemento neutro e tal que $a * e = e * a = a$ cualquiera que sea a, su imagen e' da $a' * e' = e' * a' = a'$ cualquiera que sea a' (ya que todo a' es imagen). La imagen del elemento neutro es neutra; en $E/\mathcal{F}$ corresponde a una clase neutra $C(e)$.

EJERCICIOS:

1. Aplicamos sobre $Z' = \{-1, 0, +1\}$ el conjunto Z de los enteros relativos. Las clases son: Z^+ que se aplica sobre 1; Z^- sobre -1, y 0 sobre 0. Demuéstrese que se trata de un homomorfismo para la multiplicación. ¿Elementos y clase neutros?

2. Aplicamos el conjunto $\{0, 1, 2, 3, \ldots\} = E$ sobre $E' = \{0, 1\}$. La clase $\{0, 2, 4, 6, \ldots\}$ se aplica sobre 0 y la $\{1, 3, 5, \ldots\}$ sobre 1. Pruébese que se trata de un homomorfismo para la multiplicación. ¿Elementos neutros?

3. Para E y E' del ejercicio anterior existe una ley $+'$ que corresponde a la ley $+$ en E por homomorfismo; tabúlese.

4. Señálense los endomorfismos y automorfismos de la estructura definida en el ejercicio 2 anterior.

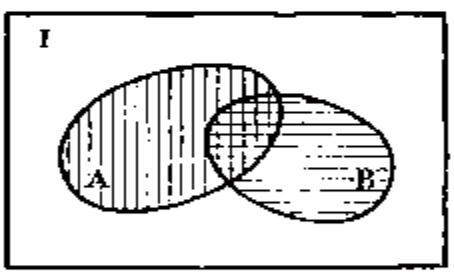
5. Se considera un cilindro de revolución (infinito) C y una de sus bases de sección recta B. Sean $\mathfrak{D}$ y $\mathfrak{d}$ los conjuntos de desplazamientos que conservan C y B, respectivamente. Hállese la f que hace de $\mathfrak{d}$ una imagen homomorfa de $\mathfrak{D}$ para la ley $\circ$. ¿Elementos y clases neutros?

ALGEBRA DE LAS PARTES DE UN REFERENCIAL

14. Una estructura muy general.—Sea I un conjunto no vacío, fijado de una vez para siempre, que llamaremos conjunto universal, o referencial. El conjunto $\mathfrak{P}(I)$ formado por todas las partes de I, incluidos ϕ e I, puede dotarse de dos leyes de composición muy naturales.

Para facilidad del lector, representaremos I por el conjunto de puntos de un rectángulo, y los subconjuntos de I mediante dominios interiores al rectángulo *(diagramas de Venn)*. De hecho, no todos los subconjuntos de un rectángulo quedan incluidos en esta categoría, que deja fuera, aparte una infinidad de otros, los subconjuntos formados por la «serpiente» infinita e infinitamente enroscada, y el árbol indefinidamente ramificado. Pero esta representación es cómoda cuando no se le pide más que servir de ayuda psicológica.

Unión.—Siendo A y B dos partes de I, llamamos *unión* de A y B y lo representamos por $A \cup B$ al conjunto de elementos que figuran en uno al menos de los conjuntos A o B.



Sobre el adjunto diagrama de Venn, $A \cup B$ comprende todo el dominio rayado, cualquiera que sea el sentido de las rayas.

Intersección.—En las mismas condiciones, se llama intersección de A y de B y se simboliza por $A \cap B$ al conjunto de elementos que figuran a la vez en A y en B.

En el diagrama de Venn anterior, $A \cap B$ comprende el dominio doblemente rayado.

Dejamos al cuidado del lector demostrar que $\cup$ y $\cap$ son dos leyes de composición internas definidas sobre $\mathfrak{P}(I)$ y que gozan de las siguientes propiedades:

1.ª idempotencia: $A \cup A = A$, $A \cap A = A$;

2.ª conmutatividad: $A \cup B = B \cup A$, $A \cap B = B \cap A$;

3.ª asociatividad: $A \cup (B \cup C) = (A \cup B) \cup C$,
$A \cap (B \cap C) = (A \cap B) \cap C$;

4.ª absorción: $A \cup (A \cap B) = A = A \cap (A \cup B)$;

5.ª distributividad: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Estas reglas pueden leerse sobre los diagramas de Venn. A título indicativo, demostremos la distributividad de la unión con relación a la intersección. Decir que x pertenece al conjunto $A \cup (B \cap C)$ significa que x pertenece a A, o que es común a B y C.

Decir que x pertenece a $(A \cup B) \cap (A \cup C)$ significa que x pertenece a la vez a A o B y a A o C; por consiguiente, x puede estar en A o ser común a B y C. Los dos conjuntos considerados coinciden.

Complementación.—Recordemos que el conjunto de elementos de I que no pertenecen a A es el complementario A′ de A.

Dejamos al lector la tarea de demostrar que la complementación está ligada a ella misma y a las otras operaciones por las siguientes propiedades:

6.ª complementariedad:

$$A \cap A' = \phi, \qquad A \cup A' = I;$$

7.ª dualidad:

$$(A \cap B)' = A' \cup B' \qquad (A \cup B)' = A' \cap B';$$

8.ª involución:

$$(A')' = A.$$

A título de ejemplo, $(A \cap B)'$ es el conjunto de los elementos que no están simultáneamente en A y en B; para ello es necesario y suficiente que estén fuera de A o fuera de B, y, por tanto, en $A' \cup B'$.

Es corriente llamar estructura algebraica a cualquier estructura que posea al menos *una* ley de composición, y *álgebra*, una estructura que posea al menos dos. Las leyes $\cup$ e $\cap$ confieren, pues, al conjunto de las partes de un referencial la categoría de *álgebra*.

EJERCICIOS:

1. Ilústrense todas las reglas de este álgebra mediante diagramas de Venn.

2. Demuéstrese que las relaciones $(A \cup C) \subset (A \cup B)$ y $(A \cap C) \subset (A \cap B)$ implican $C \subset B$.

3. Fijar dos partes A y B de I es determinar, incluyendo I y ϕ, 16 partes de I que se pide enumerar y simbolizar por medio de los signos A, B, A′, B′, $\cap$ y $\cup$.

4. Diferentes objetos O_1, O_2, O_3, ... arrojan sombra sobre un plano. Exprésense en términos de unión e intersección las reglas que determinan, a partir de las sombras de cada objeto, la formación de la sombra total y de la parte iluminada.

5. En el conjunto de las partes de I componemos A y B, quitando de su unión, $A \cup B$, su intersección, $A \cap B$. El resultado podría escribirse $(A \cup B) - (A \cap B)$; representémoslo por $A * B$.

Diagrama de Venn para $A * B$. Demuéstrese que esta ley es asociativa. Dése de $A_1 * A_2 * A_3$ una escritura simétrica en términos $\cup$ e $\cap$, y generalícese a n conjuntos.

15. Algebra de Boole.—Además de las operaciones binarias $\cup$ e $\cap$, el álgebra de las partes de un referencial admite una *operación unitaria:* la complementación. Constituye un *álgebra* de Boole[1] *concreta.*

De manera general, se denomina *álgebra de Boole* una estructura algebraica que admite tres operaciones $\cup$, $\cap$ y $'$, unidas por las propiedades 1.ª a 7.ª de la sección precedente, consideradas en este caso como *axiomas.* Se ha demostrado (M. H. STONE, 1936) que toda álgebra de Boole es isomorfa con un álgebra de Boole *concreta.*

Todo o nada.—Cuando I comprende un solo elemento, $\mathfrak{P}(I)$ comprende únicamente dos elementos I y ϕ, que representaremos más cómodamente por 1 y 0, y que podemos llamar, respectivamente, *todo* y *nada.* El álgebra de Boole de este «todo o nada» es fundamental en el estudio de las redes eléctricas en las calculadoras. Se reduce a las tres tablas:

$\cup$	1	0
1	1	1
0	1	0

$\cap$	1	0
1	1	0
0	0	0

x	x'
1	0
0	1

Se llaman *variables booleanas* aquellas variables que pueden tomar uno u otro de los valores—que se excluyen—como todo o nada, 1 ó 0, sí o no, abierto o cerrado.

EJERCICIOS:

1. En el álgebra booleana de todo o nada, se introduce la operación

$$x \oplus y = (x \cap y') \cup (x' \cap y).$$

Tabúlese.

2. Escríbanse las tablas de adición y de multiplicación del sistema binario (que emplea solo las cifras 1 y 0). Descríbase en términos booleanos

[1] George BOOLE, matemático y lógico inglés (1815-1864).

(consúltese el ejercicio anterior) la ley de formación en estas tablas del guarismo de las unidades; del guarismo de las unidades de segundo orden (colóquese un 0 si es preciso).

3. Tabúlese la función $C=(y \cap z) \cup (z \cap x) \cup (x \cap y)$: 8 casos.

4. Una álgebra de Boole concreta admite una relación de orden parcial, la relación de inclusión $A \subseteq B$. Defínase esta relación utilizando los signos $\cup$ o $\cap$. ¿Qué consecuencia se deduce para un álgebra abstracta?

5. Demuéstrese que los axiomas y primeras propiedades de un álgebra de Boole no cambian si permutamos A y A', B y B', etc., I y ϕ; después $\cup$ y $\cap$, y después $\subset$ y $\supset$. Interpretación concreta.

EJERCICIOS SOBRE EL CAPITULO II

2-1. Un conjunto $E=\{a, b, c, \ldots\}$ está dotado de una ley de composición $*$ y de una relación de orden $\leqslant$ que verifican los siguientes axiomas:

(1) $\qquad a * b \leqslant a;$

(2) $\qquad a * b \leqslant b;$

(3) $\qquad x \leqslant a \quad$ y $\quad x \leqslant b \quad$ implican $\quad x \leqslant a * b.$

A) Demuéstrese que en estas condiciones se verifica:

(4) $\qquad a * b \leqslant b * a;$

(5) $\qquad a \leqslant a * a;$

(6) $\qquad a \leqslant b \quad$ implica $\quad c * a \leqslant c * b.$

B) Pruébese que (1) y (4) y (5) y (6) implican (2) y (3).

C) Demuéstrese que $*$ es asociativa.

2-2. Sea una ley de composición no asociativa sobre un conjunto E. Llamamos $P(n)$ al número de compuestos diferentes que es posible formar con n elementos tomados en un orden determinado

a) Escribimos por convenio $P(1)=1$. Calcúlense directamente $P(2)$, $P(3)$, $P(4)$.

b) Demuéstrese que $P(n)$ verifica la fórmula de recurrencia

$$P(n)=P(n-1)P(1)+P(n-2)P(2)+\ldots \\ +P(n-k)P(k)+\ldots+P(1)P(n-1).$$

c) Para calcular $P(n)$ consideramos la serie entera [1],

$$S(x)=P(1)x+P(2)x^2+\ldots+P(n)x^n+\ldots$$

Fórmese la serie $[S(x)]^2$ y compárese con $S(x)$. Dedúzcase de ello una expresión sencilla de la suma $S(x)$ y hállense por desarrollo en serie $P(1)$, $P(2)$, ..., $P(n)$. Compruébese que

$$P(n)=\frac{(2n-2)!}{n!(n-1)!}$$

[1] Esta cuestión, que hace intervenir la teoría de las series enteras, sale del dominio estricto del álgebra.

2-3. Llamamos semirretículo todo conjunto dotado de una relación de orden parcial tal que dos elementos cualesquiera a y b admiten siempre un mínimo mayorante común, que denotaremos por $a \smile b$.

a) Demuéstrese que

$$a \smile a = a, \quad a \smile b = b \smile a, \quad (a \smile b) \smile c = a \smile (b \smile c).$$

b) Se denomina *retículo* un semirretículo tal que dos elementos cualesquiera admiten siempre un máximo minorante común. Denotaremos este por $a \frown b$; demuéstrense las relaciones análogas a las del *a)*.

c) Pruébese que $a \frown b = b$ implica $a \smile b = a$.

Demuéstrese que

$$a \frown (a \smile b) = a$$

y

$$a \smile (a \frown b) = a.$$

d) Demuéstrese que el conjunto de las partes de un referencial forma un retículo para la inclusión $\subseteq$ y que es necesario reemplazar $\frown$ por $\cap$ y $\smile$ por $\cup$. Pruébese que I es el elemento máximo absoluto y ϕ el mínimo absoluto.

e) Demuéstrese que 1, 2, 6, 7, 21, 42, forman un retículo para la relación de orden $|$, siendo el mínimo mayorante común el mínimo común múltiplo y el máximo minorante común el máximo común divisor.

f) N es un retículo para $|$. ¿Cuáles son los elementos que siguen inmediatamente a 1?

CAPITULO III

ENUMERACIONES DIVERSAS

COMBINATORIA

1. Permutaciones.—Dado un conjunto finito E, compuesto por n elementos, se llama *permutación* de E a todo conjunto totalmente ordenado según un orden estricto y compuesto de los mismos elementos que E.

Si E_4 es un conjunto formado por cuatro signos tipográficos $E_4 = \{*, \circ, \alpha, q\}$, los conjuntos ordenados definidos por

$$* < q < \alpha < 0, \qquad 0 < \alpha < * < q, \qquad q < \alpha < * < 0$$

son tres permutaciones de E_4. Dar una permutación de E equivale a fijar una ordenación de sus n elementos; toda permutación de E corresponde, pues, a una biyección de E sobre $\{1, 2, 3, \ldots, n\}$, y recíprocamente.

Existen n maneras de elegir un elemento que se aplique sobre el 1; hecha esta elección, tendremos $n-1$ maneras de elegir el que se aplique sobre el 2, después $n-2$ formas para tomar el que debe aplicarse sobre 3, etc. Existen, pues,

$$n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1$$

permutaciones de un conjunto compuesto de n elementos. Este número se representa por $n!$, que se lee *factorial de n*.

EJERCICIOS:

1. Tabúlese $n!$ para $1 \leqslant n \leqslant 10$.

2. Escríbanse todas las palabras de cuatro letras distintas que tienen las mismas letras que CIRO.

2. Variaciones.—Dado un conjunto finito E compuesto de n elementos se llama *variación* de orden p de E a todo subconjunto totalmente ordenado compuesto de p elementos de E; observemos que $p \leqslant n$.

$$* < 0 < q, \qquad * < q < 0, \qquad q < \alpha < *$$

son variaciones de orden 3 para el conjunto considerado más arriba.

Fijar una variación de orden p de E es, por consiguiente, hacer corresponder a cada uno de los números 1, 2, ..., p un elemento de E, aplicándose dos números distintos sobre dos elementos diferentes de E; la aplicación de $(1, 2, \ldots, p)$ en E debe ser inyectiva. Recíprocamente, a toda aplicación inyectiva corresponde una variación.

La imagen de 1 puede elegirse entre n elementos, la imagen de 2 entre $n-1$, ..., la imagen de p entre $n-(p-1)$. Existen, pues,

$$n(n-1)(n-2)\ldots(n-p+1)$$

variaciones de orden p para un conjunto de n elementos. Emplearemos la notación A_n^p con un subíndice y un supraíndice que no hay que confundir con un exponente.

EJERCICIOS:

1. Hágase para A_n^n la comprobación que se impone.

2. Escríbanse todas las palabras formales de 3 letras distintas, tomadas de la palabra CARIE.

3. En un plano se dan los n puntos A, B, C, ..., L. ¿Cuántos polígonos de n lados existen que tengan estos n puntos como vértices? ¿Cuántos polígonos de p lados tienen p de estos puntos como vértices?

Respuestas:

$$\frac{1}{2}(n-1)! \qquad \text{y} \qquad \frac{1}{2p}A_n^p.$$

3. Combinaciones.—Dado un conjunto finito E compuesto de n elementos, se llama *combinación* de orden p de E a todo subconjunto de E que contiene p elementos; observemos que $p \leqslant n$. Un subconjunto tal no está ordenado, lo que lo distingue de las variaciones de orden p.

De lo anterior, resulta que a cada combinación de orden p corresponden $p!$ variaciones, obtenidas considerando todos los órdenes posibles de los p elementos. Existen, pues,

$$\frac{n(n-1)(n-2)\ldots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}$$

combinaciones de orden p para un conjunto compuesto de n elementos. Representaremos este número por C_n^p y también por $\binom{n}{p}$.

EJERCICIOS:

1. Háganse, en los casos particulares, las comprobaciones que se imponen.

2. Escríbanse, respetando el orden alfabético, todas las combinaciones de orden 3 y todas las combinaciones de orden 2 del conjunto $E=\{a, b, c, d, e\}$.

3. Utilizando la noción de complementario sobre el conjunto E, demuéstrese que $C_n^p = C_n^{n-p}$. Compruébese mediante la fórmula.

4. Se ha escrito el conjunto de todas las combinaciones de orden p de un conjunto E compuesto de n signos tipográficos $E=\{a, b, c, \ldots, *, !\}$. Contando de dos modos diferentes el número total de signos escritos, demuéstrese que $pC_n^p = nC_{n-1}^{p-1}$. Dedúzcase de ello otra demostración de la fórmula que da C_n^p.

5. Hacemos una partición del conjunto anterior, por una parte, en las combinaciones que contienen la a, y, por otra, en las que no la contienen. Dedúzcase de ello que

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$

6. Se trazan n rectas en un plano, secantes dos a dos, de manera que no pasen tres por un mismo punto.

a) ¿Cuántos triángulos se forman?

b) ¿Cuántos cuadriláteros? (Comiéncese considerando que cuatro rectas determinan tres cuadriláteros: uno convexo, otro cóncavo no estrellado, y otro estrellado.)

c) Y, ¿cuántos polígonos de n lados?

7. Sea un polígono convexo de n lados; en él trazamos las diagonales (entendiendo como tales los *segmentos*). ¿Cuántos puntos de intersección? ¿Cuántos, lados y diagonales, determinan triángulos? Se distinguirán los triángulos en 4 conjuntos según que tengan 3, 2, 1 o ningún vértice confundido con un vértice del polígono.

	0	1	2	3	4	5	6	7	
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6	4	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮ .

8. Si indicamos con n la fila y con p la columna de cada uno de los números T_n^p que forman el cuadro triangular adjunto, demuéstrese que coinciden con los C_n^p con la ley $T_n^p = T_{n-1}^{p-1} + T_{n-1}^p$ (véase ejercicio 5).

4. Permutaciones con repetición.—Sea una colección[1] de n objetos en la que α ejemplares son del objeto tipo a, β ejemplares del objeto tipo b, ..., λ ejemplares del objeto tipo l; tenemos:

$$n = \alpha + \beta + \ldots + \lambda.$$

[1] Evitamos emplear la palabra *conjunto* para subrayar el hecho de que los elementos de un conjunto son discernibles.

Numerando $a_1, a_2, \ldots, a_\alpha$ los ejemplares del objeto a; $b_1, b_2, \ldots, b_\beta$ los del objeto b, etc., formamos un conjunto E de n elementos distintos. Este admite $n!$ permutaciones. Convenimos en que dos permutaciones son equivalentes si cada elemento de una de ellas pertenece al mismo tipo que el elemento que ocupa el mismo lugar en la otra. Evidentemente, se trata de una relación de equivalencia sobre el conjunto de las permutaciones de E. Todas las permutaciones de una clase se obtienen a partir de una de entre ellas por permutación de los índices de cada uno de los objetos $a, b, \ldots, l$. Tal clase comprende, en consecuencia, $\alpha!\ \beta!\ \ldots\ \lambda!$ permutaciones. Este número es independiente de la clase considerada, de donde resulta que el número de las clases es

$$\frac{n!}{\alpha!\ \beta!\ \ldots\ \lambda!}$$

Una clase se caracteriza por la colocación de los objetos de tipo determinado, independientemente de toda numeración, y se escribe, por ejemplo, *abccl ... b*. Constituye una *permutación con repetición* de los n objetos que componen la colección.

EJERCICIOS:

1. Aplíquense las consideraciones precedentes a las letras de la palabra *mississipi*, y a las de la palabra *football*.

2. Dedúzcase de la fórmula dada un teorema de aritmética.

5. Variaciones con repetición.—Sea un conjunto finito E cuyos n elementos se llaman *letras*. Una palabra formal de p letras, es decir, una sucesión ordenada de letras que se pueden repetir se denomina también *variación con repetición* de orden p de E.

La elección del elemento que ha de figurar en primera, segunda, etc., posición es siempre n, luego se ve que hay n^p variaciones de este tipo.

EJERCICIOS:

1. Redúzcase la determinación del número de las partes de un conjunto finito a un problema de variaciones con repetición para el conjunto $\{0, 1\}$.

2. Interprétese el problema de las variaciones con repetición como un problema de numeración, en el que los elementos de E sean *cifras*.

¿Cuántos números de siete cifras como máximo se pueden escribir en el sistema de base 12?

6. Combinaciones con repetición.—Consideremos un ejemplo: las K_4^6 combinaciones con repetición de orden 4 de las seis letras $\{a, b, c,$

d, e, f} pueden escribirse en el orden alfabético y ordenarse en el lexicográfico:

aaaa, aaab, ..., aaaf, aabb, ..., aaff, abbb, ..., efff, ffff.

Para estar bien seguros de poder formar una cualquiera de entre ellas con los cartones de un *juego de letras*, sería necesario disponer de cuatro ejemplares de cada una de ellas. De hecho, podemos contentarnos con un ejemplar solo, a condición de añadir a estos seis cartones otros nueve que contengan: tres el número 1; tres, el 2, y otros tres, el 3. Se puede, en efecto, aplicar biunívocamente el conjunto de las K_6^4 combinaciones con repetición sobre el conjunto de las C_{6+3}^4 combinaciones ordinarias de {*a, b, c, d, e, f*, 1, 2, 3} por medio de un código que el lector debe establecer a partir de los ejemplos que siguen:

$a\ a\ a\ a \longleftrightarrow a\ 1\ 2\ 3$ $\qquad$ $a\ a\ a\ c \longleftrightarrow a\ c\ 1\ 2$

$a\ a\ a\ b \longleftrightarrow a\ b\ 1\ 2$ $\qquad$ $a\ a\ c\ c \longleftrightarrow a\ c\ 1\ 3$

$a\ a\ b\ b \longleftrightarrow a\ b\ 1\ 3$ $\qquad$ $a\ c\ c\ c \longleftrightarrow a\ c\ 2\ 3$

$a\ b\ b\ b \longleftrightarrow a\ b\ 2\ 3$ $\qquad$ etc.,

en los que los tipos de repetición están codificados con ayuda de las cifras. De manera general, se ve que $K_n^p = C_{n+p-1}^p$

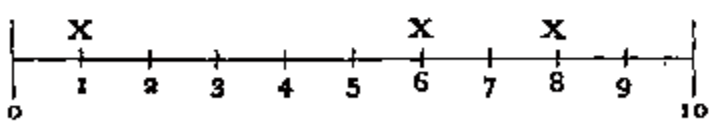
EJERCICIOS:

1. Del merengue al pastelillo relleno de crema, pasando por el pastel al ron relleno de pasas de Corinto, una pastelería expone 10 clases de pasteles. ¿De cuántas maneras podemos llenar una caja de 6?

2. Si es n un entero positivo, ¿cuántas soluciones tiene la ecuación $x_1+x_2+\ldots+x_m=n$ en los siguientes casos?

a) En números enteros estrictamente positivos.

Se podrá ver que, para $n=10$, p. ej., con $m=4$, el problema se transforma en el de hallar de cuántas maneras es posible disponer $4-1=3$ señales sobre las 9 graduaciones intermedias de un segmento que comprende 10 intervalos. El siguiente esquema, p. ej., lo simboliza



$1+5+2+2=10.$

b) En números enteros positivos o nulos.

Ahora es posible la repetición, y haciendo $y_i=x_i+1$ se reduce al caso precedente. Dedúzcase de este modo la fórmula que da K_n^p.

ENUMERACION Y ALGEBRA DE LOS CONJUNTOS

7. Unión de dos conjuntos finitos.—Siendo el referencial I finito, consideremos dos subconjuntos A y B. La unión $A \cup B$ comprende todo elemento que figura al menos en uno de los conjuntos A o B. Al contar los elementos de $A \cup B$ convenimos en no repetir los de la intersección. Si designamos por N(X) el número de elementos de una parte x de I, tendremos:

$$N(A \cup B) = N(A) + N(B) - N(A \cap B).$$

La fórmula se reduce a $N(A) + N(B)$ para dos partes disjuntas.

Caso de varios conjuntos.—Para tres conjuntos, se tiene:

$$\begin{aligned} N(A \cup B \cup C) &= N[(A \cup B) \cup C] = N(A \cup B) + N(C) - N[(A \cup B) \cap C] = \\ &= N(A) + N(B) - N(A \cap B) + N(C) - N[(A \cap C) \cup (B \cap C)] = \\ &= N(A) + N(B) + N(C) - N(A \cap B) - N(B \cap C) - N(C \cap A) + \\ &\qquad + N(A \cap B \cap C). \end{aligned}$$

Este tipo de fórmulas se extiende a un número cualquiera de conjuntos.

EJERCICIOS:

1. Consideramos los enteros n tales que $1 \leqslant n < 1000$.
¿Cuántos números existen que no son divisibles ni por 3, ni por 7? ¿Cuántos que no son divisibles ni por 5, ni por 7, ni por 11?

2. Se llama *función aditiva de conjuntos* una función f que toma sus argumentos en $\mathfrak{P}(I)$ y valores numéricos, y tal que si $A \cap B = \phi$, $f(A \cup B) = f(A) + f(B)$.

a) Demuéstrese que $f(\phi) = 0$, hallando el valor de $f(A \cup \phi)$.

b) Demuéstrese que $f(A \cup B) = f(A) + f(B) - f(A \cap B)$; se escribirá

$$A = (A \cap B) \cup (A \cap B')$$

y las fórmulas análogas.

3. Se llama función *medida* toda función aditiva que no puede tomar más que valores positivos o nulos.

a) Hállense ejemplos tales como pesos, precios, ...

Si $f(I) = 1$, f es una *probabilidad*.

b) Demuéstrese que, para una medida,

$$A \subseteq B \qquad \text{implica} \qquad f(A) \leqslant f(B).$$

FORMULAS SUMATORIAS. APLICACIONES

8. Algunas fórmulas sumatorias.—Los problemas de enumeración conducen con frecuencia a expresiones de la forma

$$1+2+3+\ldots+(n-1)+n,$$
$$1+3+5+\ldots+(2p-1),$$
$$1+4+9+\ldots+n^2,$$

que podemos calcular mediante diversos artificios. Un método general para hallar

$$f(1)+f(2)+\ldots+f(n)$$

consiste en buscar una función $F(n)$ tal que

$$f(n)=F(n)-F(n-1).$$

Escribiendo las igualdades análogas y sumando miembro a miembro, después de simplificar se obtiene:

$$f(1)+f(2)+\ldots+f(n)=F(n)-F(0).$$

Le será útil al lector el pequeño formulario que sigue, cuya comprobación (por el procedimiento de la función F) constituye un buen ejercicio de cálculo.

FORMULARIO.

$$1+2+3+\ldots\ldots\ldots+n=\frac{1}{2}n(n+1)$$

$$1+3+5+\ldots\ldots\ldots+(2p-1)=p^2$$

$$1+4+9+\ldots\ldots\ldots+n^2=\frac{1}{6}n(n+1)(2n+1)$$

$$1+9+25+\ldots\ldots\ldots+(2p-1)^2=\frac{1}{3}p(4p^2-1)$$

$$1\cdot2+2\cdot3+3\cdot4+\ldots+(n-1)n+n(n+1)=\frac{1}{3}n(n+1)(n+2)$$

$$1\cdot2\cdot3+2\cdot3\cdot4+\ldots+(n-1)n(n+1)+n(n+1)(n+2)=$$
$$=\frac{1}{4}n(n+1)(n+2)(n+3)$$

$$1^3+2^3+\ldots\ldots\ldots\ldots+n^3=\frac{1}{4}n^2(n+1)^2$$

$$1^3+3^3+\ldots\ldots\ldots\ldots+(2p-1)^3=p^2(2p^2-1)$$

EJERCICIO:

Demuéstrese que

$$[1\times 2\ldots\times r]+[2\times 3\times\ldots r\times(r+1)]+\ldots$$

$$+[n(n+1)\ldots(n+r-1)]=\frac{1}{r+1}\,n(n+1)\ldots(n+r).$$

9. Un problema a título de ejemplo.—Se trazan en el plano euclidiano K rectas secantes dos a dos, de manera que tres no concurran en un mismo punto. ¿En cuántas regiones queda dividido el plano?

Supongamos trazadas K rectas y sean N(K) las regiones. Dibujemos la $(K+1)^{\text{ésima}}$ recta, que cortará en un punto a cada una de las anteriores; estos K puntos determinan sobre la última recta introducida una partición en K+1 segmentos o semirrectas, cada uno de los cuales (o cada una de las semirrectas) se asocia con la aparición de una y una sola (hágase la figura y detállese el razonamiento) nueva región.

$$N(K+1)=N(K)+K+1.$$

De donde la tabla

$$\begin{aligned}
&K=0, && N(1)=2,\\
&K=1, && N(2)=2+2,\\
&K=2, && N(3)=2+2+3,\\
&K=n-1, && N(n)=2+2+3+\ldots+n,
\end{aligned}$$

$$N(n)=1+(1+2+3+\ldots+n)=$$

$$=1+\frac{n(n+1)}{2}=\frac{n^2+n+2}{2}$$

EJERCICIOS:

1. Problema análogo al anterior para n circunferencias secantes dos a dos.

2. Problema análogo, en el espacio, para n planos (utilícese el resultado de geometría plana).

3. Sea un polígono convexo de n lados, del que se trazan las diagonales, de las cuales no concurren tres en un mismo punto. ¿Cuántas regiones hay en el interior del polígono?

Resp.:

$$f(n+1)=f(n)+\frac{(n-2)(n-1)n}{6}+n-1,$$

de donde,

$$f(n)=\frac{(n-1)(n-2)(n^2-3n+12)}{24}$$

4. Si $2n$ personas participan en un torneo de tenis, ¿de cuántas maneras puede organizarse la primera vuelta (cada persona debe jugar una sola vez)?

EJERCICIOS SOBRE EL CAPITULO III

3-1. *Un problema de Cayley.*—¿Cuántos polígonos convexos de k lados existen cuyos vértices coinciden con los vértices y los lados con las *diagonales* de un polígono convexo de n lados dado? *Resp.:* $n/k\, C^{k-1}_{n-k-1}$.

3-2. Doce personas que trabajan en cierta comisión tienen acceso a documentos secretos encerrados en una caja fuerte. ¿Cuántas cerraduras diferentes es necesario prever como mínimo y cuántas llaves hay que dar a cada persona para que 5 de ellas reunidas no pueden abrir nunca la caja y sí lo puedan hacer siempre 6?

3-3. Sea un cuadrado ABCD de lado n. Lo dividimos en n^2 cuadrados de lados iguales a la unidad mediante rectas paralelas a AB y AD. Referimos los puntos del plano tomando A como origen y los ejes Ax y Ay sobre AB y AD, respectivamente.

1.° Unimos A con un vértice o nudo cualquiera $M(x, y)$ del cuadriculado mediante un contorno poligonal cuyos lados sean todos lados de la cuadrícula y que verifique la siguiente propiedad: si se recorre el contorno de A a M la abscisa y la ordenada no decrecen nunca.

Calcúlese el número de contornos que unen A y M. Encuéntrese de esta manera la clásica fórmula relativa a las combinaciones:

$$C^r_p = C^r_{p-1} + C^{r-1}_{p-1}.$$

En particular, determínese el número de tales poligonales que unen A y C.

2.° Afectamos a cada vértice del cuadriculado un coeficiente igual al número de poligonales que le unen al punto A. Calcúlese la suma de los coeficientes atribuidos a los vértices situados:

a) sobre una misma paralela a la diagonal BD, que corta al segmento AB;

b) en el interior del triángulo ABD, o sobre su contorno;

c) sobre una misma paralela a uno de los lados del cuadrado;

d) en el interior del cuadrado ABCD o sobre su contorno.

3.° Se designa por Σ el área plana limitada por una poligonal del tipo anterior que une A con C, así como por AB y BC. Calcúlese la suma de las áreas Σ obtenidas para todos los contornos posibles.

4.° Consideramos el segmento MM' del cuadriculado que une los puntos $M(x, y)$ y $M'(x+1, y)$, $0 \leqslant x \leqslant n-1$ y $0 \leqslant y \leqslant n$. Calcúlese el número Π^y_x de poligonales que unen A y C y que contengan todas el segmento MM'.

Calcúlense las sumas

$$\Pi^0_x + \Pi^1_x + \Pi^2_x + \ldots + \Pi^n_x,$$
$$\Pi^y_0 + \Pi^y_1 + \Pi^y_2 + \ldots + \Pi^y_{n-1};$$

se demostrará que son respectivamente independientes de x y de y; indíquese qué relación presenta esta segunda suma con la cuestión de la sección 3 anterior.

CAPITULO IV

ESTRUCTURA DE GRUPO

LOS AXIOMAS DE LA ESTRUCTURA DE GRUPO

1. Definición.—Se llama grupo[1] todo conjunto no vacío, G, dotado de una ley de composición interna, asociativa, con elemento neutro, y tal que cada uno de sus elementos admite un simétrico.

En otras palabras: si las variables x, y, z, ..., x', ... tienen G como dominio de definición, la estructura de grupo se rige por los cuatro axiomas:

1.°	$(\forall x)(\forall y)(\exists x * y)$	:	$(x * y) \in G$;
2.°	$(\forall x)(\forall y)(\forall z)$	:	$(x * y) * z = x * (y * z)$;
3.°	$(\exists e)(e \in G)(\forall x)$	:	$e * x = x * e = x$;
4.°	$(\forall x)(\exists x')$	:	$x * x' = x' * x = e$.

Los grupos para los que la ley $*$ es conmutativa se denominan *abelianos*[2].

Consecuencias inmediatas de los axiomas.—De las generalidades expuestas en los primeros capítulos, resulta inmediatamente:

TEOREMA 1.—*El compuesto de n elementos de un grupo, considerados en cierto orden, no depende de las agrupaciones hechas a lo largo del cálculo* (teorema de asociatividad).

TEOREMA 2.—*El elemento neutro es único.*

TEOREMA 3.—*El elemento simétrico es único.*

Isomorfismo.—Sean dos conjuntos G y G' entre los que existe una biyección regular respecto de las leyes $*$ y $*'$, respectivamente definidas sobre G y G'. Si la ley $*$ da a G una estructura de grupo (es decir, verifica los axiomas 2.°, 3.° y 4.° anteriores), sabemos (véase pág. 37)

[1] Concepto y nombre debidos a Evaristo GALOIS, genial matemático francés nacido en 1811, muerto en duelo a la edad de 21 años.

[2] Niels Henrik ABEL (1802-1829), gran matemático noruego.

que la ley $*'$ también los verifica; G' es, por consiguiente, un grupo *isomorfo* a G.

EJERCICIOS:

1. Demuéstrese, sin consultar el libro, que el elemento neutro es único, y que el elemento inverso es único.

2. ¿Por qué no es un grupo para la unión el álgebra de Boole, de las partes de un referencial? ¿Y para la intersección?

3. $E=\{1\}$ dotado de la multiplicación y $F=\{0\}$ dotado de la adición son dos grupos isomorfos.

4. $G=\{-1, +1\}$ dotado de la multiplicación es un grupo.

5. *Axiomas débiles.*—Demuéstrese que los axiomas 3.° y 4.° se pueden reemplazar por los axiomas más débiles siguientes:

3.° bis: existe (al menos) un elemento neutro a la derecha, e, tal que $ae=a$ para todo $a \in G$;

4.° bis: existe (al menos) un elemento inverso a la derecha, a', tal que $aa'=e$.

Método: Considérese x tal que $a'x=e$; demuéstrese primero que a' es también inverso a la izquierda, y, después, que es único.

2. Otras propiedades elementales.—Los axiomas de estructura de grupo implican directamente las siguientes propiedades:

TEOREMA 4.—*Todos los elementos de un grupo son regulares.*

Si $a * b = a * c$, multiplicando a la izquierda por el simétrico a' de a los dos miembros de esta igualdad, obtenemos:

$$a' * (a * b) = a' * (a * c)$$

de donde, por asociatividad,

$$(a' * a) * b = (a' * a) * c$$

o sea,

$$e * b = e * c;$$

por consiguiente,

$$b = c,$$

resultado al que llegaríamos también partiendo de $b * a = c * a$.

TEOREMA 5.—*En un grupo las ecuaciones* $a * x = b$ *e* $y * a = b$ *admiten, cada una, solución única.*

En efecto,

$$a * x = b \iff a' * a * x = a' * b \iff x = a' * b$$

e

$$y * a = b \iff y * a * a' = b * a' \iff y = b * a'.$$

Para las notaciones a emplear en los diferentes casos, se consultará el cuadro que figura a continuación.

3. Vocabulario y notaciones.—El siguiente cuadro da la nomenclatura en uso.

	Para una ley representada por ·, o por ausencia de signo, y eventualmente por ∘, o por *	Para una ley representada por + (exclusivamente reservada a ciertos grupos abelianos).
El grupo es calificado de	multiplicativo	aditivo
El compuesto se llama	producto	suma
El elemento neutro toma el nombre de	unidad, elemento unidad; identidad en el caso ∘	cero, elemento nulo
El elemento neutro se representa por	e, I, I, I, ..., según los casos	0, 0, O, ..., según los casos.
El elemento simétrico toma el nombre de	inverso	opuesto
El elemento simétrico de a se representa por	a^{-1}	$-a$
Las soluciones de $a * x = b$, $y * a = b$ se representan por	$x = a^{-1}b$ $y = ba^{-1}$ y b/a si es abeliano	$x = y = b + (-a)$ o $b - a$
$a * a * \ldots * a$ (p factores) se representa por	$a \cdot a \cdot \ldots \cdot a = a^p$	$a + a + a + \ldots + a = pa$

EJERCICIOS:

1. Pruébese que, sobre un grupo (multiplicativo) cualquiera, la ecuación $axbcx = abx$ admite solución única. Determínese esta.

2. Sobre un grupo multiplicativo cualquiera hállese *una* solución de la ecuación $xax = bba^{-1}$.

3. Sobre un grupo abeliano multiplicativo, resuélvase la ecuación $xabxc = bxa$.

4. Si a y b son dos elementos cualesquiera de un grupo multiplicativo, demuéstrese que existe un elemento x, llamado *conmutador a la derecha*, tal que $abx = ba$, y un elemento y, *conmutador a la izquierda*, tal que $yab = ba$. ¿Cómo están ligados los conmutadores de a y b a los de b y a? ¿Y de b^{-1} y a^{-1}? ¿Y de a^{-1} y b^{-1}?

5. Sobre el conjunto $Q \times Q$ de pares ordenados de racionales (a, b) tales que $a \neq 0$, definimos la ley $(a, b) * (a', b') = (aa', ba' + b')$. Demuéstrese que esta ley confiere al conjunto estructura de grupo.

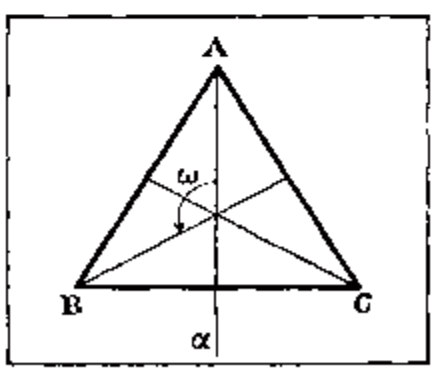
4. Ejemplos de grupos. Grupos finitos.—El conjunto N de los enteros naturales, completado con el cero, no constituye un grupo ni para la adición ni para la multiplicación, las dos leyes fundamentales de la aritmética. Existe el elemento neutro: cero y uno, respectivamente, *pero falta el simétrico*, tanto para una como para la otra.

Las dos ampliaciones clásicas de la noción de número, fracciones y enteros relativos, tienen por meta prolongar el conjunto N para obtener grupos. Ilustremos la noción de grupo con algunos ejemplos, con objeto de familiarizar al lector con esta estructura tan difundida en matemáticas.

En geometría, los movimientos directos o desplazamientos—que son aplicaciones particulares del espacio sobre sí mismo, con conservación de las longitudes de los segmentos y del sentido de los triedros orientados—forman un grupo para la ley asociativa $\circ$. Las nociones de desplazamiento neutro (identidad) y de desplazamiento inverso son bien conocidas. De manera general, las biyecciones de un conjunto en sí mismo constituyen un grupo para la ley $\circ$; la asociatividad, así como la existencia de un elemento neutro y de un simétrico han sido demostradas en el capítulo II.

Pero si un conjunto está dotado de estructura, toda biyección de este conjunto, regular respecto de la estructura considerada (véase página 37), se llama *automorfismo* de dicha estructura. Los automorfismos de una estructura forman un grupo.

Grupos finitos.—Los grupos mencionados anteriormente se componían de infinitos elementos. Se llama grupo *finito* cualquier grupo que contiene un número finito de elementos; este número es el *orden* del grupo. En los ejercicios relativos a la sección 1 se han citado algunos ejemplos; vamos a ver otro, tomado de la geometría.



Dado en el espacio, en magnitud y posición, un triángulo equilátero ABC, consideremos las operaciones geométricas llamadas *isometrías* que aplican este triángulo sobre sí mismo (véase figura). Estas son en número de seis, a saber: la operación neutra e; la rota-

ción ω, alrededor del eje del triángulo, de 1/3 de vuelta en sentido directo; la rotación θ, inversa de la anterior; la simetría α, cuyo eje es la bisectriz del ángulo de vértice A, y las simetrías análogas β y γ. Según la composición habitual de las aplicaciones, realizar dos operaciones en un orden determinado equivale a efectuar una sola, su producto, que denotaremos con el signo $\circ$, respetando la escritura de derecha a izquierda. A título de ejemplo hacemos el producto de θ por α:

A viene sobre C por θ, después C viene sobre B por α: $A \longrightarrow B$,
B viene sobre A por θ, después A viene sobre A por α: $B \longrightarrow A$,
C viene sobre B por θ, después B viene sobre C por α: $C \longrightarrow C$;

$\circ$	e	ω	θ	α	β	γ
e	e	ω	θ	α	β	γ
ω	ω	θ	e	γ	α	β
θ	θ	e	ω	β	γ	α
α	α	β	γ	e	ω	θ
β	β	γ	α	θ	e	ω
γ	γ	α	β	ω	θ	e

por consiguiente,

$$\gamma = \alpha \circ \theta.$$

Todos los resultados análogos se consignan en la tabla de la ley $\circ$ para el conjunto $\mathcal{G} = \{e, \omega, \theta, \alpha, \beta, \gamma\}$. La ley $\circ$ es asociativa, el conjunto es cerrado para $\circ$, e es el elemento neutro y cada operación posee su inversa: $\mathcal{G}$ es un grupo no conmutativo de orden 6.

EJERCICIOS:

1. En un grupo finito de orden par, además del elemento neutro, existe un número impar de elementos que son sus propios inversos.

2. Escríbase la tabla de $\mathcal{G}$ (véase apartado anterior) cambiando el orden de las columnas de tal manera que todos los e estén sobre la diagonal; en estas condiciones, ¿qué axioma «queda destacado»?

3. Estúdiese el grupo de las isometrías de un rectángulo. Hállese su orden; reconózcase que es abeliano; hállense sus automorfismos y el grupo que estos forman. Compárese este último con el grupo $\mathcal{G}$.

4. Sobre la recta numérica completada por un punto ∞ se consideran las cuatro funciones f_1, f_2, f_3, f_4 tales que

$$x \xrightarrow{f_1} x, \qquad x \xrightarrow{f_2} \frac{1}{x}, \qquad x \xrightarrow{f_3} -x, \qquad x \xrightarrow{f_4} -\frac{1}{x}$$

Demuéstrese que forman grupo para la ley $\circ$. Compárese este con el grupo del rectángulo.

5. En las condiciones del ejercicio precedente, estúdiese el grupo de las seis funciones

$$x \to x, \qquad x \to \frac{1}{x}, \qquad x \to 1-x,$$

$$x \to \frac{1}{1-x}, \qquad x \to \frac{x-1}{x}, \qquad x \to \frac{x}{x-1}$$

y compárese con $\mathcal{J}$.

6. Estúdiese el grupo de las isometrías de un cuadrado; tabla.

7. En la tabla de composición de un grupo finito, demuéstrese que cada elemento figura una vez, y solo una, en cada línea y en cada columna.

8. Se llama *cuadrado latino* un cuadrado cuyas n líneas y n columnas contienen todas los mismos elementos, $a_1, a_2, \ldots, a_n$; cada elemento figura una sola vez por línea y por columna. Demuéstrese que la condición necesaria para que la tabla de un grupo sea un cuadrado latino no es suficiente. Utilícese el cuadrado latino definido en la página 29, ejercicio núm. 3.

9. Demuéstrese que para un grupo *finito* los axiomas 3.° y 4.° (pág. 54) pueden reemplazarse por:

3.° triplicado: G contiene un número finito de elemento ·

4.° triplicado: todos los elementos de G son regulares.

5. Subgrupos.—Sucede a veces que una parte H de un grupo G forma ella misma un grupo; se dice entonces que H es un subgrupo de G. Todo grupo admite dos subgrupos impropios: él mismo y el conjunto reducido al solo elemento neutro.

Hallemos las condiciones necesarias y suficientes para que H, con $H \subset G$, sea un subgrupo. Según el axioma 1.°, H debe ser cerrado para la ley de G. El axioma 2.° se cumple siempre. El axioma 3.° exige que H contenga el elemento neutro. Finalmente, según el axioma 4.°, H debe contener el inverso de cada uno de sus elementos.

De hecho, estas condiciones se condensan en una sola.

TEOREMA.—*Para que la parte H de G sea un subgrupo es necesario y suficiente que para todo par (x, y) de elementos de H, H contenga también* $x * y^{-1}$.

Es decir, $(\forall x \in H)(\forall y \in H) : (x * y^{-1}) \in H$.

En efecto, con (x, x), H, contiene a $x * x^{-1} = e$. Con (e, x), H contiene a $e * x^{-1} = x^{-1}$ y, finalmente, con x e y, H contiene a x e y^{-1}; por consiguiente, $x * (y^{-1})^{-1} = x * y$.

Los múltiplos de 2, de 3, etc., forman subgrupos del conjunto de los enteros relativos; las potencias de un número racional forman un subgrupo del grupo de los racionales. En general, si $a \in G$, el subconjunto $\{\ldots, a^{-2}, a^{-1}, a^0, a^1, a^2, \ldots\}$ es un subgrupo de G.

En geometría, las traslaciones forman un subgrupo del grupo de los desplazamientos. Las rotaciones que conservan un punto fijo O forman otro subgrupo.

En el ejemplo del grupo $\mathcal{G}$ del triángulo equilátero, tenemos los siguientes subgrupos (los tres últimos son isomorfos):

$\circ$	e	ω	θ
e	e	ω	θ
ω	ω	θ	e
θ	θ	e	ω

$\circ$	e	α
e	e	α
α	α	e

$\circ$	e	β
e	e	β
β	β	e

$\circ$	e	γ
e	e	γ
γ	γ	e

El primero conserva las caras del triángulo (considerado como una placa); los otros, las cambian, pero conservan, respectivamente, A, B y C. En general, aquellas de las operaciones de un grupo de transformaciones que dejan «alguna cosa» fija forman un subgrupo.

EJERCICIOS:

1. Determínense los subgrupos del grupo de las isometrías de un rectángulo.

2. Estúdiense los subgrupos del grupo del cuadrado (uno de ellos es isomorfo al precedente).

3. La intersección de dos subgrupos es siempre un subgrupo. ¿Cabe decir otro tanto de la unión?

4. En el conjunto de los pares de racionales (a, b) dotados de la ley $(a, b) * (a', b') = (aa', ba' + b')$, los pares $(1, b)$ forman un subgrupo.

5. En un grupo G, el conjunto de los elementos permutables con todos los otros $[(\forall x)\, ax = xa]$ constituye el centro de G. Demuéstrese que es un subgrupo. ¿Cómo encontrarlo en la tabla de G cuando este es finito?

6. Reflexiónese sobre el grupo de las transformaciones que aplican sobre sí mismo un reticulado infinito de malla cuadrada. ¿Cómo engendrar este grupo? ¿Cuántas traslaciones es necesario dar? ¿Cuántas simetrías? Búsquense subgrupos finitos, infinitos.

GRUPOS DE SUSTITUCIONES

6. Definición.—Se llama *sustitución* una aplicación de un conjunto finito sobre sí mismo. Necesariamente, una sustitución es biunívoca, porque si dos elementos (al menos) tuvieran la misma imagen habría (al menos) un elemento sin antecedente.

Una sustitución s transforma una permutación π del conjunto E de

n elementos sobre el que actúa en otra permutación π'. Por consiguiente, las sustituciones son operadores externos para el conjunto Π de las $n!$ permutaciones de E, y escribiremos

$$s \cdot \pi = \pi' \qquad \text{o} \qquad s = \left(\frac{\pi'}{\pi}\right)$$

Hemos llegado de esta forma a designar s por una *fracción simbólica,* que se obtiene escribiendo como denominador sobre una línea los n elementos de E (o los signos que los representan biunívocamente) ordenados en el orden de π y, encima, sus imágenes respectivas, ordenadas en el orden de π'. Si la permutación denominador es la π_1, la permutación numerador se convierte en π'_1, tal que se conserven los pares de elementos en correspondencia. Tenemos, pues,

$$s = \left(\frac{\pi'}{\pi}\right) = \left(\frac{\pi'_1}{\pi_1}\right).$$

Para el conjunto de E formado por los cinco signos tipográficos:

$$\mathrm{E} = \{*,\ 0,\ \alpha,\ q,\ ?\},$$

tendremos de esta forma cierta sustitución s denotada indiferentemente por

$$s = \left(\frac{\alpha\ ?\ q\ 0\ *}{*\ 0\ \alpha\ q\ ?}\right) = \left(\frac{0\ q\ \alpha\ *\ ?}{q\ \alpha\ *\ ?\ 0}\right) = \left(\frac{*\ \alpha\ ?\ 0\ q}{?\ *\ 0\ q\ \alpha}\right) = \ldots$$

Si fijamos arbitrariamente la permutación denominador, existen $n!$ numeradores posibles y, por consiguiente, $\mathrm{N} = n!$ sustituciones; sean $s_1, s_2, \ldots, s_N$.

El grupo simétrico.—Las sustituciones que se efectúan sobre un conjunto dado E se componen, como las aplicaciones (de las que son un caso particular), según la ley asociativa representada por $\circ$. Existe una sustitución neutra que aplica todo elemento sobre sí mismo, y toda sustitución $s = \left(\frac{\pi'}{\pi}\right)$ admite una inversa representada por la fracción simbólica invertida $\left(\frac{\pi}{\pi'}\right)$. Estas sustituciones forman, pues, un grupo de orden $n!$ A partir de otro conjunto de n elementos obtendríamos evidentemente un grupo isomorfo con el primero; todos los grupos «concretos» análogos son isomorfos a un mismo grupo abstracto, el *grupo simétrico* $\mathfrak{S}_n$ (letra gótica). Demostremos que $\mathfrak{S}_3$ es isomorfo al

grupo del triángulo equilátero (véase pág. 57). En efecto, toda transformación que conserva el triángulo establece una sustitución entre los vértices A, B, C: a cada uno de ellos le hace corresponder aquel cuyo lugar va a ocupar. Así, p. ej., a la rotación ω le corresponde la sustitución $\begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$, etc. Las definiciones de las leyes de composición sobre estos dos grupos hacen de esta correspondencia biunívoca un isomorfismo.

EJERCICIOS:

1. Sea $E=\{1, 2, 3, 4\}$. Demuéstrese que las cuatro sustituciones

$$e=\begin{pmatrix} 1&2&3&4 \\ 1&2&3&4 \end{pmatrix}, \quad r=\begin{pmatrix} 2&1&4&3 \\ 1&2&3&4 \end{pmatrix}, \quad s=\begin{pmatrix} 3&4&1&2 \\ 1&2&3&4 \end{pmatrix}, \quad t=\begin{pmatrix} 4&3&2&1 \\ 1&2&3&4 \end{pmatrix}$$

forman un subgrupo de $\mathfrak{S}_4$. Compárese este subgrupo con el grupo del rectángulo.

2. TEOREMA DE CAYLEY.—Todo grupo finito de orden n, G, puede representarse como un grupo de sustituciones hechas entre sus elementos; por consiguiente, como un subgrupo de $\mathfrak{S}_n$.

Método: sea $G=\{e, a, b, \ldots, l\}$; a $x \in G$ se asociará la sustitución s tal que, para todo $a \in G$, tengamos $s(a)=xa$.

7. Transposición.—Se llama *transposición* una sustitución que aplica cada elemento sobre sí mismo, a excepción de dos de entre ellos a y b; la imagen de a es, por tanto, b, y recíprocamente.

De ello resulta que el producto $t \circ t=t^2$, de una transposición t por ella misma, es la transformación idéntica i; tenemos, pues, $t=t^{-1}$.

TEOREMA.—*Toda sustitución puede descomponerse en producto de transposiciones.*

El teorema es cierto para dos elementos A y B, porque el grupo $\mathfrak{S}_2$ comprende la transposición t que cambia A y B y la sustitución idéntica $i=t^2$. Razonando por recurrencia, la suponemos cierta para $\mathfrak{S}_{n-1}$. Sea entonces s una sustitución de $\mathfrak{S}_n$ hecha entre n letras A, B, ..., K, L. Si existe en s un elemento A que es su propia imagen, la restricción de s al conjunto de los restantes elementos es una sustitución s_1 del grupo $\mathfrak{S}_{n-1}$, que puede realizarse como producto de transposiciones. Asociando a toda transposición de $\mathfrak{S}_{n-1}$ una transposición de $\mathfrak{S}_n$ que conserve A, vemos que también s es un producto de transposiciones.

Si s no conserva ninguno de los n elementos, sea D la imagen de A, y t la transposición que cambia A y D. La sustitución $s'=t \circ s$ (que

denotamos por ts) conserva A y puede representarse como producto de transposiciones; de donde,

$$ts = t_k \dots t_2 t_1 \qquad \text{y} \qquad s = tt_k \dots t_2 t_1.$$

El teorema es cierto para las sustituciones del grupo $\mathfrak{S}_n$.

EJERCICIOS:

1. Descompónganse en producto de transposiciones las sustituciones

$$\left(\frac{2\,4\,5\,3\,1}{1\,2\,3\,4\,5}\right) \qquad \text{y} \qquad \left(\frac{2\,3\,1\,5\,6\,4}{1\,2\,3\,4\,5\,6}\right).$$

2. Efectúese[1]

$$\left(\frac{2\,1}{1\,2}\right) \circ \left(\frac{6\,5}{5\,6}\right) \circ \left(\frac{3\,2}{2\,3}\right) \circ \left(\frac{5\,4}{4\,5}\right) \circ \left(\frac{4\,1}{1\,4}\right) \circ \left(\frac{6\,4}{4\,6}\right),$$

y hágase después la descomposición en otro producto más sencillo de transposiciones.

3. Se llama *sustitución circular* o *ciclo* la sustitución que se obtendría disponiendo los elementos en los vértices de un polígono regular y atribuyendo como imagen de cada elemento el que ocupa el vértice siguiente (en un sentido dado). Demuéstrese que $\left(\frac{5\,4\,6\,2\,3\,1}{1\,5\,4\,6\,2\,3}\right)$ es un ciclo.

4. Escríbase

$$s = \left(\frac{5\,6\,2\,1\,3\,7\,4\,9\,8}{1\,2\,3\,4\,5\,6\,7\,8\,9}\right)$$

en la forma $\left(\frac{5\,3\,2\,6\,7\,4\,1\,9\,8}{1\,5\,3\,2\,6\,7\,4\,8\,9}\right)$ y reconózcase que s es producto de dos ciclos (véase ejercicio 3). Demuéstrese que toda sustitución puede descomponerse en un producto de ciclos hechos sobre conjuntos disjuntos.

5. Demuéstrese que el orden de un ciclo es igual al número de elementos sobre el que opera. El orden del producto de dos ciclos s y t, hechos sobre dos conjuntos disjuntos que comprenden p y q elementos, es el m.c.m. de p y q.

8. Paridad de una sustitución.—Se comprueba fácilmente sobre un ejemplo que una sustitución dada puede descomponerse de varias maneras en producto de transposiciones; vamos a precisar una pro-

[1] Las sustituciones consideradas se hacen sobre las seis cifras 1, 2, 3, 4, 5, 6. Denotamos por $\left(\frac{1\,2}{2\,1}\right)$ la transposición que cambia los dos primeros elementos, dejando fijos los otros.

piedad interesante de las posibles descomposiciones, pero examinemos primero el caso de la sustitución idéntica.

TEOREMA 1.—*Toda descomposición de la sustitución idéntica en producto de transposiciones entraña un número par de factores.*

Razonemos sobre n elementos y precisemos el número de ellos que se conservan por un producto de transposiciones. Una transposición t_1 conserva $n-2$ elementos y cambia los otros dos; un producto t_2t_1 de dos transposiciones conduce a la sustitución idéntica si $t_1=t_2$; si no, conserva $n-3$ ó $n-4$ elementos según que los cambios se hagan sobre tres o cuatro elementos *distintos*. Pasemos a un producto $t_3t_2t_1$; se demuestra fácilmente que conserva $n-2$ elementos a lo más y $n-6$ como mínimo, pero es imposible conservar los n elementos. Añadiendo otros factores t_4, t_5, ..., nos damos cuenta que sucede lo mismo para un producto cualquiera que comprenda un número impar de transposiciones: el máximo número de elementos conservados es $n-2$. Un tal producto nunca dará la sustitución idéntica.

TEOREMA 2.—*Toda descomposición de una sustitución dada en un producto de transposiciones entraña un número de factores cuya paridad está bien determinada.*

Sean $s=t_k \dots t_2t_1$ y $s=t'_l \dots t'_2t'_1$ dos descomposiciones de la sustitución s, comprendiendo una de ellas k factores, y la otra, l. Ya que una transposición es su propia inversa, tendremos $s^{-1}=t'_1t'_2 \dots t'_l$ y, por consiguiente,

$$i=ss^{-1}=t_k \dots t_2t_1t'_1t'_2 \dots t'_l.$$

Según el teorema 1 el número $k+l$ es par y, por tanto, k y l son de la misma paridad.

Consecuencias.—Podemos clasificar las sustituciones en dos clases: pares e impares. En particular, toda transposición t es evidentemente impar.

Si t es una transposición fija, la relación $s'=st$ (o $s=s't$) asocia a toda sustitución par s una sustitución impar s', y recíprocamente. Hay, pues, el mismo número $n!/2$ de sustituciones pares que impares.

El grupo alternado.—Se reconoce inmediatamente que el producto de dos sustituciones pares es una sustitución par, y que también lo es la inversa de toda sustitución par. (Por el contrario, el producto de dos impares es par.) Las sustituciones pares que se efectúan sobre un con-

junto de n elementos forman así un subgrupo de $\mathfrak{S}_n$ llamado grupo *alternado*, que se designa por $\mathfrak{A}_n$ (letra gótica).

Signatura.—Se llama signatura de una sustitución s al número

$$\epsilon_s = \begin{cases} +1 \text{ si } s \text{ es par,} \\ -1 \text{ si } s \text{ es impar.} \end{cases}$$

EJERCICIOS:

1. Demuéstrese que el grupo de las operaciones geométricas que aplican un tetraedro regular sobre sí mismo es isomorfo con $\mathfrak{S}_4$. Distínganse los movimientos directos de los inversos. ¿Cuál es el subgrupo correspondiente a $\mathfrak{A}_4$?

2. Estúdiese en función de n la paridad de la sustitución que asocia $(n, n-1, \ldots, 2, 1)$ a $(1, 2, \ldots, n)$.

ESTUDIO DE LOS SUBGRUPOS. AUTOMORFISMO. HOMOMORFISMO DE GRUPOS

9. Partición estable de un grupo.—Para designar un subconjunto no vacío de un grupo dado G se emplea la palabra *complejo*. Si $A = \{a, a_1, a_2, \ldots\}$ es un complejo, se denota por $x * A$ el complejo

$$\{x * a, x * a_1, x * a_2, \ldots\}$$

que se llama el trasladado de A a la izquierda por x. La aplicación de A sobre $x * A$ es biunívoca y estos dos complejos encierran el mismo número de elementos (si este número es finito); en general, tienen la misma potencia. De la misma forma se define el trasladado a la derecha $A * x$. Hallemos si puede existir una partición de G en complejos A, B, ..., disjuntos, tal que, cualquiera que sea $x \in G$, los complejos $x * A$, $x * B$, ... realicen la misma partición de G, salvo el orden, evidentemente. Una partición tal se denomina estable a la izquierda. Los complejos A, B, ... tienen necesariamente la misma potencia; además, si $x * A$ e $y * B$ poseen un elemento común, coincidirán, ya que son dos complejos de la partición inicial, disjunta.

TEOREMA 1.—*En una partición estable, el complejo* E *que contiene al elemento neutro es un subgrupo.*

Sea $E = \{e, e_1, e_2, \ldots\}$ el complejo que contiene al elemento neutro e. Puesto que $e_i * E$ contiene a $e_i * e = e_i$, debe coincidir con E. En con-

secuencia, $e_i * E$ es biunívocamente aplicable sobre E; a todo elemento $e_i * e_j$ corresponde $e_k \in E$, y a todo elemento e_j un e_l tal que $e_l * e_j = e$. Estas dos condiciones son necesarias y suficientes para que E constituya un subgrupo de G.

TEOREMA 2.—*Los complejos* A, B, C, E, ... *de una partición estable a la izquierda de un grupo* G *se deducen de un subgrupo* E *de* G *por traslaciones a la izquierda.*

Si A contiene el elemento a, como $a = a * e$, vemos que A y $a * E$ poseen un elemento común; por consiguiente, coinciden. Todas las clases se forman, pues, de la manera indicada.

OBSERVACIONES.—1.ª Se demuestra fácilmente la recíproca: a partir de todo subgrupo $E \subset G$ se obtiene una partición estable a la izquierda considerando los complejos $x * E$, que son distintos.

Los complejos $E * x$ distintos proporcionan una partición estable a la derecha.

2.ª Para un grupo abeliano las nociones se confunden.

3.ª La condición para x e y de pertenecer a un mismo complejo es una relación de equivalencia 8. Se tiene: $x 8 y$ si, y solo si, $y^{-1} * x \in E$, como se comprueba fácilmente.

En el caso de que A esté confundido con E, se encuentra de nuevo la condición para obtener un subgrupo.

TEOREMA.—*En un grupo finito el orden de todo subgrupo divide al orden del grupo.*

Si n es el orden del subgrupo E y p el número de clases de la partición estable asociada, el orden de G es igual a np, ya que las clases son equipotentes.

El número p se llama *índice* del subgrupo.

Subgrupo distinguido o invariante.—Se denomina subgrupo *distinguido* o *invariante* un subgrupo que da una partición estable a la derecha y a la izquierda.

Si N es un tal subgrupo $x * N$, por ser en la partición una clase compuesta a la izquierda, será también una clase compuesta a la derecha; por consiguiente, $x * N = N * x$ (tienen común x, y, por tanto, coinciden).

En un grupo abeliano todo subgrupo es necesariamente distinguido; en un grupo no conmutativo no sucede así; se demostrará como ejercicio que para el grupo del triángulo equilátero, $\{e, \alpha\}$ no es distinguido, mientras que $\{e, \omega, \theta\}$ sí lo es.

EJERCICIOS:

1. Fórmense las clases compuestas a la derecha y a la izquierda de $\{e, \alpha\}$ en $\mathcal{G}$ (véase pág. 58) y compruébese que son diferentes.

2. Clases a la izquierda y después a la derecha de $\{e, \omega, \theta\}$ en $\mathcal{G}$; compruébese que realizan la misma partición. ¿Cuál es el índice de este subgrupo?

3. Se marcan sobre un eje infinito los puntos M_n de abscisas enteras. Sea G el grupo de las transformaciones geométricas (traslaciones y simetrías) que aplican esta configuración sobre ella misma. Estúdiense las clases:
   a) para el subgrupo engendrado por una simetría;
   b) para el subgrupo de las traslaciones.

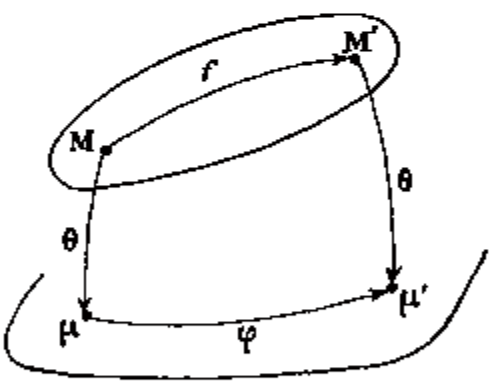
4. Se llama centro de un grupo al conjunto de los elementos que conmutan con todos los elementos del grupo. Demuéstrese que el centro es un subgrupo distinguido.

5. La intersección de dos subgrupos distinguidos es un subgrupo distinguido.

6. Demuéstrese que $\mathfrak{M}_4$ contiene un subgrupo distinguido isomorfo al grupo del rectángulo.

10. Subgrupos conjugados. Automorfismos internos.—Recordemos en primer lugar algunas nociones empleadas en geometría. Sea el punto M' que deriva del punto M por la transformación f; si sometemos el conjunto a una transformación θ, M da μ y M' da μ'. Pasamos entonces de μ a μ' por la cadena

$$\mu \xrightarrow{\theta^{-1}} M \xrightarrow{f} M' \xrightarrow{\theta} \mu'$$



y la transformación $\varphi = \theta \circ f \circ \theta^{-1}$ se llama *transmutada* de f por θ.

En teoría de grupos diremos que a_1 *es conjugado* de a (abreviadamente, $a_1 \mathcal{C} a$) si existe en el grupo G un elemento x tal que a_1 sea transmutado de a por x; dicho de otra manera:

$$a_1 \mathcal{C} a \iff (\exists x) : a_1 = x * a * x^{-1}.$$

Demostremos que $\mathcal{C}$ define una relación de equivalencia sobre G:

1.° $a = e * a * e^{-1}$;

2.° $a_1 = x * a * x^{-1} \Longrightarrow a = x^{-1} * a * x = (x^{-1}) * a * (x^{-1})^{-1}$;

3.° $a_1 = x * a * x^{-1}$ y $a_2 = y * a_1 * y^{-1} \Longrightarrow a_2 = y * (x * a * x^{-1}) * y^{-1}$
$= (y * x) * a * (x^{-1} * y^{-1}) = (y * x) * a * (y * x)^{-1}$.

$\mathcal{C}$ permite, pues, descomponer G en clases disjuntas que, contrariamente a las clases del apartado anterior, no son equipotentes. En particular, ya que $(\forall x)\ x * e * x^{-1} = e$, el elemento neutro e no es conjugado más que de sí mismo.

Automorfismo interno.—Si tomamos en G un elemento fijo s, la aplicación σ y su inversa σ^{-1}, determinadas por

$$a \xrightarrow{\sigma} a_1 = s * a * s^{-1}; \qquad a_1 \xrightarrow{\sigma^{-1}} a = s^{-1} * a_1 * s$$

definen un isomorfismo de G consigo mismo; es decir, un automorfismo que se califica de *interno*. En efecto, se ve fácilmente que σ da como imagen del compuesto de dos elementos el compuesto de las imágenes:

$$a_1 * b_1 = (s^{-1} * a * s) * (s^{-1} * b * s) = s^{-1} * (a * b) * s.$$

El lector demostrará igualmente que la imagen de la inversa es la inversa de la imagen.

TEOREMA.—*Los transmutados de los elementos de un subgrupo* H, *mediante un automorfismo interno* σ *del elemento* s, *forman un subgrupo isomorfo con* H.

La demostración es muy sencilla y se propone como ejercicio al lector. Este nuevo subgrupo recibe el nombre de *conjugado* de H.

En el grupo de los desplazamientos del espacio euclidiano, si R y R_1 son los subgrupos de las rotaciones que dejan fijos dos puntos dados, O y O_1, respectivamente, y si T designa la traslación de vector $\mathbf{OO_1}$, se tiene $R_1 = T \circ R \circ T^{-1}$.

TEOREMA.—*Una condición necesaria y suficiente para que un subgrupo sea distinguido es que coincida con todos sus conjugados.*

Sea N un subgrupo invariante; para todo x, se tiene $x * N = N * x$, de donde $N = x * N * x$, y recíprocamente.

En un grupo abeliano

$$a_1 = x * a * x^{-1} \Longrightarrow a_1 = x * x^{-1} * a = a;$$

el único isomorfismo interno es neutro, y todos los subgrupos son invariantes.

EJERCICIOS:

1. Se transmuta una inversión f por otra inversión θ; demuéstrese que en la transmutada φ dos puntos y sus homólogos son siempre concíclicos. Dedúzcase de ello que φ es una inversión.

2. Sean f_1 y f_2 dos inversiones positivas; estúdiese el producto $f_2 \circ f_1$, transmutándolo por una inversión θ que da para el conjunto de las dos circunferencias de inversión una figura canónica (dos rectas secantes o paralelas, o dos circunferencias concéntricas).

3. En $\mathcal{G}$, $\{e, \alpha\}$, $\{e, \beta\}$, $\{e, \gamma\}$ son conjugados; $\{e, \omega, \theta\}$ coincide con sus conjugados.

4. Hállense, geométricamente, subgrupos conjugados en el grupo de las transformaciones que aplican sobre sí mismo un tetraedro regular.

5. En un grupo abeliano, todo elemento no es conjugado más que de sí mismo. *Aplicación:* ¿Cuál es la transmutada de una traslación T por otra traslación $\mathfrak{T}$?

6. Descompóngase el grupo $\mathcal{G}$ del triángulo en clases de elementos conjugados.

11. Equivalencia regular.—La estructura de grupo es una de las más simples del álgebra. Prosigamos su estudio y supongamos que sobre el conjunto G existe, además de la ley $*$, una relación de equivalencia, $\mathcal{E}$. Esta hipótesis tiene particular interés cuando la ley $*$ es estable respecto de $\mathcal{E}$ (véase pág. 24); vamos a precisar las propiedades del conjunto cociente G/$\mathcal{E}$. En lugar de afirmar que la ley es estable respecto de $\mathcal{E}$, muchas veces es más cómodo decir que la equivalencia es *regular* respecto de la ley.

TEOREMA 1.—*Si una equivalencia* $\mathcal{E}$ *es regular respecto de la ley de composición de un grupo, la ley inducida sobre el conjunto cociente es también una ley de grupo.*

La demostración es inmediata en virtud del teorema de la página 28 (completado con el ejercicio 7 de la pág. 29). Hemos obtenido, pues, un nuevo grupo G', llamado *grupo factor* (también *grupo cociente*). Caractericemos sus elementos, es decir, las clases de equivalencia respecto de $\mathcal{E}$; en particular, sea E la clase unidad formada por el conjunto $\{e, e_1, e_2, \ldots\}$ de los elementos equivalentes a e.

TEOREMA 2.—*En toda partición de un grupo por una equivalencia regular, la clase unidad,* E, *es un subgrupo invariante de* G. *Las otras clases son los diferentes complejos que resultan por traslación a la derecha (o a la izquierda) de* E.

En efecto, si e_1 y e_2 son dos elementos cualesquiera de E, las relaciones $e_1 \mathcal{E} e$, $e_2 \mathcal{E} e$ implican

$$e_1 * e_2 \mathcal{E} e * e \mathcal{E} e; \qquad \text{por tanto,} \qquad (e_1 * e_2) \in \mathrm{E}.$$

Sea e_1^{-1} el inverso de e_1; en la igualdad $e_1 * e_1^{-1} = e$ reemplacemos e_1 por su equivalente e; el nuevo resultado, equivalente a e, es $e_3 \in \mathrm{E}$,

de donde $e * e_i^{-1} = e_3$, o $e_i^{-1} = e_3 \in E$. El subconjunto E es, por consiguiente, un subgrupo de G.

Finalmente, sea A la clase de equivalencia asociada con un elemento a de G, y sea x un elemento cualquiera de A. El elemento $x * a^{-1}$ es equivalente a $a * a^{-1} = e$; este es, pues, un $e_i \in E$, de donde

$$x * a^{-1} = e_i, \qquad x = e_i * a.$$

Todo $x \in A$ pertenece así al complejo Ea y, recíprocamente, Ea está formado por elementos equivalentes al a; en consecuencia, $A = Ea$.

Considerando $a^{-1} * x$, veríamos también que $A = aE$, de donde se deducen las dos consecuencias siguientes:

α) E es un subgrupo invariante;

β) las clases de equivalencia son los complejos asociados a este subgrupo. El grupo factor G' se designa muchas veces con la notación G/E. El teorema precedente admite un recíproco de fácil demostración:

Recíproco.—*Todo subgrupo invariante* E *de un grupo* G *determina una equivalencia regular respecto de la ley de* G *y un grupo factor* G/E.

Para un grupo abeliano G, cualquier subgrupo puede desempeñar este papel. Examinemos el caso importante del grupo aditivo de los enteros relativos Z.

Clases residuales en Z.—Sobre Z la adición es una ley de grupo conmutativo. Si m es un entero cualquiera se comprueba fácilmente que la relación $a - b = km$ entre dos enteros a y b (con $k \in Z$) es una relación de equivalencia regular respecto de la adición. El subgrupo invariante asociado, M, está formado por el conjunto de los múltiplos de m, y las otras clases, en número de $m - 1$, son: la clase $M + 1$ (formada por los números $a = km + 1$), la clase $M + 2$, ..., la clase $M + m - 1$. Todas son equipotentes y constituyen un grupo finito de m elementos para la ley inducida; se denominan *clases residuales* (o *clases de restos) módulo m*. La ley de adición para estas clases es la misma que la de Z, considerando como iguales los números que se diferencian en un múltiplo de m. El lector encontrará en el capítulo siguiente la tabla de adición del grupo para $m = 6$.

12. Homomorfismo para los grupos.—Sea $G = \{e, a, b, \ldots\}$ un grupo para le ley $*$, y $G' = \{a', b', \ldots\}$, un dominio homomorfo dotado de una ley $*'$. Existe, pues, una suprayección h de G en G' compatible

con las leyes de composición; es decir, que

$$a \overset{h}{\longrightarrow} a', \qquad b \overset{h}{\longrightarrow} b' \Longrightarrow a * b \overset{h}{\longrightarrow} a' *' b'.$$

El siguiente teorema precisa una propiedad fundamental de G'.

TEOREMA.—*Una imagen homomorfa de un grupo es un grupo.*

Sabemos que h determina una equivalencia $\mathcal{H}$ sobre G, estando formadas las clases por elementos cuya imagen por h es la misma. Comprobemos que se trata de una equivalencia regular respecto de la ley $*$. Si a y a_1 tienen la misma imagen a', y es b' la imagen común de b y b_1, la compatibilidad muestra que

$$a * b \overset{h}{\longrightarrow} a' *' b' \qquad \text{y} \qquad a_1 * b_1 \overset{h}{\longrightarrow} a' *' b';$$

los elementos $a * b$ y $a_1 * b_1$ son, por consiguiente, equivalentes respecto de $\mathcal{H}$.

En virtud del teorema de la página 39, el conjunto cociente $G/\mathcal{H}$ dotado de la ley inducida es isomorfo a G' dotado de $*'$. Pero, como acabamos de ver, $G/\mathcal{H}$ es un grupo y G' un grupo isomorfo.

Núcleo.—Los elementos de G cuya imagen es el elemento neutro e' de G' (es decir, los elementos equivalentes a e respecto de $\mathcal{H}$) constituyen un subgrupo invariante de G llamado *núcleo* del homomorfismo.

EJEMPLOS. 1. Sea G el conjunto de los vectores libres del espacio ordinario y P un plano fijo. Hagamos corresponder a todo vector su proyección ortogonal sobre P. Esta proyección es un homomorfismo, ya que la proyección de la suma da la suma de las proyecciones. G' es el grupo de los vectores libres paralelos a P. El núcleo está constituido por el grupo de los vectores perpendiculares a P.

2. El grupo del triángulo equilátero comprende tres rotaciones, e, ω, θ, que conservan el signo de los ángulos orientados, y tres simetrías, α, β, γ, que multiplican este signo por (-1). Si aplicamos el grupo sobre el grupo multiplicativo $\{-1, +1\}$ se obtiene un homomorfismo de núcleo $\{e, \omega, \theta\}$. Compruébese el isomorfismo del grupo factor.

13. Grupos monógenos. Grupos cíclicos.—Un grupo G se llama *monógeno* cuando es engendrado por los compuestos de un solo elemento a:

$$G = \{\ldots a^{-2}, a^{-1}, a^0, a, a^2, a^3, \ldots\};$$

un grupo monógeno finito es un grupo *cíclico*.

El conjunto Z dotado de la adición es monógeno, porque es engendrado por el solo elemento 1, cuyo simétrico (en este caso el opuesto) es -1. En el grupo $\mathcal{G}$ del triángulo equilátero, el subgrupo formado por las rotaciones planas, $e=\theta^0=\theta^3$, $\theta=\theta^{-1}$, $\omega=\theta^2$, es un ejemplo de grupo cíclico de tercer orden.

Sea G el grupo monógeno engendrado por el elemento a; la aplicación h de Z sobre G definida por

$$p \overset{h}{\longrightarrow} a^p$$

respeta las leyes de grupo. Es un homomorfismo, cuya realización depende de que G sea infinito o finito.

En el primer caso todos los a^p son distintos y h es un isomorfismo, lo que conduce al siguiente teorema:

TEOREMA.—*Todo grupo monógeno infinito es isomorfo al grupo aditivo de los enteros relativos* Z.

En el segundo caso, no todos los a^p son distintos. Si G admite m elementos, las potencias $a^0=e$, a, a^2, ..., a^{m-1} lo son necesariamente y la repetición comienza solo con a^m, que debe dar de nuevo e, como se demuestra fácilmente por reducción al absurdo. La aplicación h es en este caso un homomorfismo, cuyo núcleo es el subgrupo de los múltiplos de m en Z. La teoría precedente nos conduce inmediatamente al resultado expresado por el siguiente teorema.

TEOREMA.—*Todo grupo cíclico* G *de orden* m *es isomorfo al grupo aditivo de las clases residuales módulo* m *en* Z.

El cálculo dentro de G se reduce de nuevo al cálculo módulo m, como para las rotaciones del m-ágono regular. Este resultado explica el calificativo de cíclico atribuido a G.

Completemos este apartado con una observación sobre los grupos finitos.

TEOREMA.—*En un grupo finito, todo elemento* a *engendra un grupo cíclico.*

En efecto, a engendra un grupo monógeno, que es finito como subgrupo de un grupo finito; por consiguiente, cíclico. El orden de este grupo cíclico se llama, abreviadamente, *orden de* a.

EJERCICIOS:

1. Demuéstrese que el orden de un elemento divide al orden del grupo.
2. Determínese el orden de cada elemento del grupo del cuadrado.

3. En un grupo cíclico el elemento a es de orden p, y el elemento b, de orden q. Hállese el orden del elemento ab.

4. ¿Bajo qué condiciones el elemento a de un grupo cíclico engendra el grupo entero? Un elemento de este tipo se llama *primitivo*. ¿Cuáles son los elementos primitivos de Z/10, de Z/12? Dése la interpretación geométrica por medio de polígonos regulares.

14. Compuesto de dos grupos.—Sean $G=\{e, a, b, c, \ldots, x, \ldots\}$ y $\Gamma=\{\epsilon, \alpha, \beta, \gamma, \ldots, \xi, \ldots\}$ dos grupos respecto de dos leyes, respectivamente denotadas por los signos $\perp$ y $\wedge$.

Formemos el producto $G\times\Gamma$ que dotaremos de la ley producto; dicho de otra manera:

$$(x, \xi) * (x_1, \xi_1) = (x \perp x_1, \xi \wedge \xi_1).$$

El conjunto $G\times\Gamma$ es así un grupo con el elemento neutro (e, ϵ), el elemento inverso (x^{-1}, ξ^{-1}), etc.

Si G y Γ son multiplicativos, se dice que $G\times\Gamma$ es su *producto directo;* si son abelianos denotados aditivamente, diremos que se ha formado su *suma directa;* en este último caso el nuevo grupo sigue siendo abeliano.

Sean $G=Z/2=\{\bar{0}, \bar{1}\}$ y $\Gamma=Z/3=\{\bar{0}, \bar{1}, \bar{2}\}$; su suma directa comprende los pares $(\bar{0}, \bar{0})$, $(\bar{0}, \bar{1})$, $(\bar{0}, \bar{2})$, $(\bar{1}, \bar{0})$, $(\bar{1}, \bar{1})$ y $(\bar{1}, \bar{2})$. Este grupo es cíclico y resulta engendrado por los múltiplos de $(\bar{1}, \bar{1})$ que se repiten con periodicidad $2\times 3=6$; es isomorfo a Z/6.

EJERCICIOS:

1. Demuéstrese que el grupo del rectángulo es isomorfo a la suma directa $(Z/2)+(Z/2)$.

2. Siendo p y q dos enteros primos entre sí, pruébese que el producto directo de un grupo cíclico de orden p por otro grupo cíclico de orden q es un grupo cíclico de orden pq. *Aplicación:* El grupo de las rotaciones planas del dodecágono.

EJERCICIOS SOBRE EL CAPITULO IV

4-1. Consideramos las funciones homográficas o bilineales,

$$x \xrightarrow{f} \frac{ax+b}{cx+d}, \qquad ad-bc\neq 0,$$

definidas sobre la recta proyectiva (es decir, sobre la recta numérica completada con un punto ∞ tal que $\frac{1}{\infty}=0$). Demuéstrese que el conjunto de estas funciones forma un grupo para la composición de las aplicaciones.

Hállense los subgrupos finitos ayudándose de los ejercicios propuestos en el apartado de los grupos finitos.

4-2. Consideremos sobre $Z \times Z$ las aplicaciones f del tipo

$$(x, y) \longrightarrow (X, Y) \quad \text{con} \quad X = ax + by, \quad Y = cx + dy; \quad ad - bc = 1.$$

Pruébese que forman un grupo que conserva el m.c.d. de los pares.

4-3. Demuéstrese que existe un solo grupo abstracto de orden 2, y un solo grupo abstracto de orden 3. Pruébese que únicamente existen dos grupos abstractos de orden 4, uno isomorfo con Z/4, el otro isomorfo con el grupo del rectángulo.

4-4. Búsquense todos los grupos finitos que tengan n elementos para $5 \leqslant n \leqslant 9$. Descompónganse, si es posible, en producto directo o suma directa. Indíquense los grupos abelianos.

4-5. Sean z_1 y z_2 dos ciclos que solo tengan un elemento común en sus dominios de definición. Demuéstrese que $z_1 \circ z_2$ es un ciclo. ¿De qué orden? Compárense $z_1 \circ z_2$ y $z_2 \circ z_1$. Utilícense los resultados obtenidos para probar que toda sustitución par puede descomponerse en un producto de ciclos, de modo que cada uno opere sobre tres elementos.

4-6. Consideramos el grupo alternado $\mathfrak{A}_n$ para $n \geqslant 4$ y suponemos que contiene un subgrupo invariante N.

a) Demuéstrese que si N contuviera un ciclo $\begin{pmatrix} b\,c\,a \\ a\,b\,c \end{pmatrix}$ de orden 3, contendría todos y coincidiría con $\mathfrak{A}_n$.

b) Pruébese que N no podría contener sustituciones de orden 3.

c) Demuéstrese que N no se podría descomponer en trasposiciones.

Concluyase de lo anterior que $\mathfrak{A}_n$, para $n \geqslant 4$, no posee subgrupo invariante (Galois).

4-7. Sea G un grupo abeliano; se dice que los elementos a de orden α, b de orden β, c de orden γ, ..., l de orden λ forman una *base* de G si todo elemento x de G puede expresarse en la forma $x = a^{A} b^{B} \ldots l^{L}$.

Demuéstrese (si se cree necesario, sobre grupos de orden poco elevado) que todo grupo abeliano finito tiene una base, y que los órdenes α, β, ..., λ de los elementos de esta son de la forma $\alpha = p_1^{K_1}$, $\beta = p_2^{K_2}$, etc., donde los números p_i son primos.

Pruébese que estos $p_i^{K_i}$ son invariantes para los cambios de base. ¡Demuéstrese que dado un conjunto $p_i^{K_i}$ existe un grupo abeliano que los admite como invariantes! (Pueden emplearse espacios vectoriales sobre cuerpos módulo p_i.)

CAPITULO V

ESTRUCTURA DE ANILLO

LOS AXIOMAS DE LA ESTRUCTURA DE ANILLO

1. Definición.—Dado un grupo abeliano A, que denotaremos aditivamente, definimos sobre A una estructura de anillo dotando a este conjunto de una segunda ley interna, llamada multiplicación, asociativa y doblemente distributiva con relación a la adición.

Los axiomas de la estructura de anillo son los siguientes:

I. Grupo abeliano aditivo.—Sean x, y, z elementos cualesquiera de A.

1.° Existe una ley que denotamos por $+$ $(\forall x)(\forall y): (x+y) \in A$.

2.° $(\forall x)(\forall y): x+y=y+x.$

3.° $(\forall x)(\forall y)(\forall z): (x+y)+z=x+(y+z).$

4.° $(\exists 0)(0 \in A)(\forall x): 0+x=x+0=x.$

5.° $(\forall x)(\exists -x): (-x)+x=x+(-x)=0.$

II. Segunda ley interna: *multiplicación.*

1.° Existe una ley denotada por $\times$, o $\cdot$, o incluso por ausencia de signo.

$(\forall x)(\forall y): \quad (xy) \in A.$

2.° $(\forall x)(\forall y)(\forall z): \quad (xy)z=x(yz).$

III. Axiomas que ligan las dos leyes.

1.° $(\forall x)(\forall y)(\forall z): \quad x(y+z)=xy+xz.$

2.° $(\forall x)(\forall y)(\forall z): \quad (y+z)x=yx+zx.$

OBSERVACIONES DIVERSAS.—Cuando la multiplicación es conmutativa, el anillo se llama conmutativo. Existen, en efecto, anillos no conmutativos, de los que el más importante es el de las matrices cuadradas.

$A-\{0\}$ no es necesariamente un grupo para la multiplicación, e

incluso puede no poseer elemento neutro para esta ley. El anillo Z estudiado anteriormente es un ejemplo de anillo con *elemento unidad*, en el que sus elementos carecen de inversos (excepción de 1 y -1). El conjunto de los números pares es otro ejemplo de anillo *sin elemento unidad*.

2. **Consecuencias de la existencia de un grupo.**—Como los elementos tienen estructura de grupo abeliano para la adición, escribamos las propiedades que de ello resultan (véanse págs. 54 y 55).

α) El elemento neutro 0 es único.

β) El opuesto de un elemento es único.

γ) El opuesto del opuesto de un elemento es igual a este elemento:

$$-(-a)=a.$$

δ) El opuesto de una suma es la suma de los opuestos de cada sumando:

$$-(a+b+c)=(-a)+(-b)+(-c).$$

ϵ) Ley de simplificación para la adición: la relación $a+c=b+c$ implica $a=b$.

ζ) La ecuación $x+b=a$ admite solución única $x=a+(-b)$. Por definición, escribimos $x=a-b$ y el elemento x recibe el nombre de diferencia entre a y b.

El convenio $a-b=a+(-b)$ reduce el cálculo de una diferencia al de una suma. En virtud de la conmutatividad de esta última operación, el resultado de una serie de adiciones y sustracciones es independiente del orden de los términos:

$$a-b-c+d=a+d-b-c=-b-c+a+d=\ldots$$

En álgebra elemental una expresión como la anterior se llama suma algebraica.

Según δ), el opuesto de una suma algebraica se obtiene haciendo la suma algebraica de los opuestos de cada término:

$$-(a-b-c+d)=-a+b+c-d.$$

De ello se deduce la regla de supresión de paréntesis utilizada en el álgebra elemental:

$$a+(-b+c-d)=a-b+c-d,$$
$$a-(-b+c-d)=a+b-c+d.$$

3. Otras propiedades.—α) Cualquiera que sea $a \in A$, se tiene que $a \times 0 = 0$. Por definición, el elemento neutro es tal que $b + 0 = b$. Multiplicando por a:

$$a(b + 0) = ab,$$

de donde

$$ab + a \times 0 = ab = ab + 0.$$

Por la ley de simplificación, nos queda $a \times 0 = 0$.

Como A no se ha supuesto conmutativo, es necesario respetar el orden de los factores. Se demostraría igualmente que $0 \times a = 0$.

β) Si en un producto se sustituye un factor por su opuesto, el producto se cambia en su opuesto. Comprobemos, p. ej., que

$$(-a)b = -(ab),$$

es decir,

$$(-a)b + ab = 0.$$

Operando sobre el primer miembro y aplicando a propiedad anterior:

$$(-a)b + ab = [(-a) + a]b = 0b = 0.$$

Se demuestra igualmente que $a(-b) = -ab$.

Para calcular $(-a)(-b)$, aplicamos dos veces la propiedad precedente:

$$(-a)(-b) = -[a(-b)] = -[(-ab)] = ab.$$

Vemos que es válida la regla de los signos, independientemente de la existencia de elementos positivos o negativos. Se aplica al producto de elementos opuestos.

γ) Combinando las leyes obtenidas, el lector comprobará que, sobre un anillo conmutativo, a toda expresión donde intervengan sumas, diferencias y productos pueden aplicarse las mismas reglas operatorias que para los números relativos. Basta, por ello, saber efectuar un producto de dos sumas algebraicas

$$A = (a - b + c)(-d + e - f).$$

Se multiplica término a término aplicando la regla β) y se efectúa la suma algebraica de los resultados obtenidos. Si el anillo no es conmutativo, será necesario respetar siempre en los productos el orden de los factores.

4. Anillo de integridad.—Sobre un anillo, la ley de simplificación es siempre válida para la adición, pero no necesariamente para la multiplicación. La igualdad

$$ac = bc$$

no implica, pues, $a = b$, incluso para $c \neq 0$. En particular, así es en el anillo de las matrices cuadradas. Indiquemos una condición equivalente a la ley de simplificación.

TEOREMA.—*La ley de simplificación para la multiplicación es válida sobre un anillo, para todo elemento no nulo, si, y solo si, el producto de dos elementos diferentes de cero es distinto de cero.*

Partiendo de $ac = bc$, con $c \neq 0$, se deduce que $ac - bc = 0$, de donde

$$(a - b)c = 0.$$

Si suponemos que el producto de elementos distintos de cero es diferente de cero, necesariamente tendremos que $a - b = 0$, o sea $a = b$, ya que $c \neq 0$. Es, por tanto, válida la ley de simplificación.

Recíprocamente, demostremos que, si puede satisfacerse la igualdad $a'b' = 0$, con $a' \neq 0$ y $b' \neq 0$, no es válida la ley de simplificación. Si a es un elemento cualquiera, tendremos:

$$ab' = ab' + a'b' = (a + a')b';$$

si la ley de simplificación fuera válida para $b' \neq 0$, se tendría $a = a + a'$, de donde $a' = 0$, contra la hipótesis.

Un anillo *conmutativo* en el que es válida la ley de simplificación se denomina *anillo de integridad*. Para un anillo que no sea de integridad existen ciertos elementos no nulos cuyo producto es nulo: los llamados *divisores de cero* (véase pág. 89).

Finalmente, observemos que encontraremos muchas veces el caso de anillos de integridad en los que la multiplicación admite elemento neutro e. Diremos que se trata de *dominios de integridad*.

Subanillos.—En todo anillo A, cualquier subconjunto que sea un subgrupo del grupo aditivo, cerrado con relación a la ley multiplicativa, posee una estructura de anillo y constituye un subanillo de A.

Así, p. ej., en el anillo de los enteros relativos estudiado anteriormente, los múltiplos de un número fijo a forman un subanillo. Es de señalar que el elemento neutro de A para la ley multiplicativa puede no pertenecer al subanillo, el cual no será, por consiguiente, un dominio de integridad.

EJERCICIOS:

1. Un grupo abeliano denotado aditivamente da un anillo si definimos en él una segunda ley $a \cdot b = 0$ cualesquiera que sean a y b. (Anillo de cuadrado nulo.)

2. Demuéstrese que los elementos 0 y e forman un anillo para las leyes:

$$\begin{array}{ll} 0+e=e+0=e, & e\cdot 0=0\cdot e=0, \\ 0+0=0, & 0\cdot 0=0, \\ e+e=0, & e\cdot e=e. \end{array}$$

Interprétese lo anterior reemplazando 0 por *par* y e por *impar*.

3. Sea A un anillo tal que $x^2=x$ para todo $x \in A$.

a) Demuéstrese, escribiendo la propiedad precedente para $x+x$, que $x=-x$ para todo $x \in A$.

b) Pruébese, escribiendo la propiedad precedente para $x+y$, que A es conmutativo.

c) Demuéstrese que la relación $xy=x$ entre dos elementos x e y de A es una relación de orden, que denotaremos por $x \mathfrak{R} y$.

d) Observando que $xy(x+y)=0$, para todo x y para todo y, demuéstrese que si A no tiene divisores de cero, o bien A contiene solo el elemento 0, o bien A no contiene más que dos elementos, 0 y x.

5. Anillo ordenado.—Un anillo A se llama *ordenado* cuando existe sobre A una relación de orden total compatible con la ley de adición. Si esta relación se denota por $\leqslant$, se tendrá, para cualquier c,

$$a \leqslant b \iff a+c \leqslant b+c.$$

En particular, $c=-a$ da $0 \leqslant b-c$. El elemento $b-c$ se llama *positivo* (o nulo para la igualdad); vamos a caracterizar en primer lugar la clase de estos elementos.

Elementos positivos.—Se supone que en el anillo A sabemos definir un subconjunto, representado por A^+, que verifica los siguientes axiomas:

a) la suma de dos elementos de A^+ pertenece a A^+;

b) el producto de dos elementos de A^+ pertenece a A^+;

c) si a es un elemento cualquiera de A, necesariamente se verifica una, y solo una, de las posibilidades siguientes:

$$a \in A^+; \qquad (-a) \in A^+; \qquad a=0.$$

El subconjunto A^+ define los elementos positivos del anillo; sus opuestos constituyen el conjunto A^- de los elementos negativos. El elemento 0 no es ni positivo ni negativo.

TEOREMA.—*El producto de dos elementos pertenecientes al mismo subconjunto* (A^+ *o* A^-) *es un elemento de* A^+. *El producto de dos elementos que pertenecen uno a* A^+, *y el otro a* A^-, *es un elemento de* A^-.

Si a y b pertenecen a A^+, sabemos que $ab \in A^+$; si a y b pertenecen a A^-, $(-a)$ y $(-b)$ pertenecen a A^+ y, según las reglas de cálculo, $ab=(-a)(-b)$ pertenece a A^+.

Finalmente, si $a \in A^+$ y $b \in A^-$, se tiene que $ab=-(a)(-b)$ y ab es el opuesto de un elemento de A^+, por lo que pertenece a A^-.

En particular, el cuadrado de un elemento cualquiera es siempre positivo; si se trata de un anillo unitario, el elemento e tal que $e=e\cdot e$ es también positivo.

En resumen, sobre uno de estos anillos obtenemos la regla de los signos tal y como se enseña en el álgebra elemental.

Orden total.— En un anillo en el que existe un subconjunto A^+ de elementos positivos, escribimos $a \leqslant b$ cuando la diferencia $b-a$ es positiva o nula.

Demostremos que queda definida de esta manera una relación de orden total compatible con la adición:

a) tenemos $a \leqslant a$;

b) si $a \leqslant b$ y $b \leqslant a$, $a=b$. En efecto, $b-a$ pertenece a $\{A^+, 0\}$ y su opuesto $a-b$ a $\{A^-, 0\}$. Pero la última desigualdad prueba que $a-b$ pertenece también a $\{A^+, 0\}$. Ya que los dos subconjuntos A^+ y A^- son disjuntos, será necesariamente $a=b$.

c) Si $a \leqslant b$ y $b \leqslant c$, se tiene que $a \leqslant c$; conclusión inmediata en virtud de la tercera propiedad de los elementos positivos.

Finalmente, la compatibilidad respecto de la adición se deduce inmediatamente de la fórmula

$$b-a=(b+c)-(a+c).$$

Respecto de la multiplicación, se tiene la siguiente propiedad:

TEOREMA.—*Si* $c \in A^+$, *la relación* $a \leqslant b$ *implica que* $ac \leqslant bc$. *Si* $c \in A^-$, *la misma relación implica que* $bc \leqslant ac$ [1].

La demostración clásica del álgebra elemental es válida en este caso.

Señalemos, por fin, que la relación $a < b$ es equivalente a la $a \leqslant b$, con $a \neq b$.

Valor absoluto.—En un anillo totalmente ordenado el valor absoluto

[1] Relación que se escribe también $ac \geqslant bc$.

de un elemento cualquiera a no nulo, simbolizado por $|a|$, es por definición el elemento positivo del par a, $-a$. Dicho valor es nulo para $a=0$. Satisface las dos propiedades siguientes:

TEOREMA.—*Para dos elementos cualesquiera a y b, se verifica siempre que*

$$|ab|=|a|\times|b|, \qquad |a+b|\leqslant |a|+|b|.$$

La primera relación es inmediata para dos elementos positivos. Si estos son negativos, se tiene $a=-|a|$ y $b=-|b|$, de donde

$$ab=(-|a|)\times(-|b|)=|a|\times|b| \qquad \text{y} \qquad ab=|ab|$$

porque ab es positivo.

Para a negativo y b positivo, $a=-|a|$, $b=|b|$, de donde $ab=-|a|\times|b|$: pero ab es negativo, y, en consecuencia, $|ab|=-ab=|a|\times|b|$.

La segunda relación se establece observando que un elemento a cualquiera verifica la doble desigualdad

$$-|a|\leqslant a\leqslant |a|.$$

De igual modo,

$$-|b|\leqslant b\leqslant |b|.$$

Sumando,

$$-(|a|+|b|)\leqslant a+b\leqslant |a|+|b|.$$

Si $(a+b)\in A^+$, $(a+b)=|a+b|$ y la desigualdad buscada resulta de la que figura a la derecha en la relación precedente.

Si $(a+b)\in A^-$, se tiene que $(a+b)=-|a+b|$ y la desigualdad buscada se deduce, después de multiplicar por $-e$, de la desigualdad de la izquierda en dicha relación.

EJERCICIOS:

1. En un anillo totalmente ordenado:

1.° Si $a\leqslant b$ y $c\leqslant d$, $a+c\leqslant b+d$. (Se pueden sumar miembro a miembro dos desigualdades del mismo sentido.)

2.° Si $c\in A^+$, la desigualdad $ac\leqslant bc$ exige $a\leqslant b$ (recíproco del segundo teorema de la página 80).

3.° La suma de dos elementos negativos es también un elemento negativo.

4.° La igualdad $x+x=0$ implica $x=0$; de igual modo, $nx=0$ implica $x=0$ para cualquier n entero natural (véase pág. 90).

5.° Para dos elementos cualesquiera a y b, $||a|-|b||\leqslant|a-b|$.

6.° La ecuación $x^2+e=0$ carece de solución.

7.° Para $c>0$, la igualdad $a=b+c$ implica $a<b$.

8.° Cualesquiera que sean a y b, se tiene que $a^2-ab+b^2>0$.

9.° La desigualdad $a \leqslant b$ implica $a^3 \leqslant b^3$.

2. Demuéstrese que la característica de un anillo ordenado es necesariamente infinita (véase pág. 90).

EL ANILLO Z DE LOS ENTEROS RELATIVOS

6. Introducción.—El lector está habituado a manejar los números relativos (positivos, negativos o nulos), pero la teoría que de ellos se hace en álgebra elemental es poco satisfactoria.

Nos limitaremos aquí al conjunto Z de los enteros relativos, los cuales forman grupo respecto de la adición; la operación inversa, la sustracción, es siempre posible. Además, ciertos elementos de este grupo aparecen como casi idénticos a los enteros naturales (números positivos), y los otros son sus opuestos (a excepción del 0). Construiremos aquí Z a partir de los enteros naturales, pero definiendo un conjunto más amplio sobre el que sea siempre posible la sustracción. Transponemos al plano teórico el método utilizado de ordinario para hacer posible una sustracción momentáneamente imposible.

Imaginemos un cajero que puede recibir o pagar dinero. Si tiene actualmente 5 ptas. en caja y debe efectuar un pago de 8 ptas. ello es imposible, pero puede pedir a otra caja el préstamo de una suma suficiente, 10 ptas., p. ej. El razona entonces sobre un par (haber, debe), en este caso (15, 10), que le permite efectuar su pago, obteniendo el par (7, 10). Si las siguientes operaciones le conducen a un haber superior al debe, podrá reembolsar, y ello le conducirá a un par de la forma $(n, 0)$. En resumen, las operaciones con los enteros (adición y sustracción) se presentan como casos particulares de operaciones sobre pares de enteros. Para precisar estas operaciones (así como la multiplicación) examinaremos en primer lugar los pares (a, b) con $a \geqslant b$ y haremos que su conjunto sea isomorfo al de los enteros. Se extiende luego la teoría a pares cualesquiera, lo que conduce a la representación clásica de los enteros relativos.

7. Representación de los enteros por pares.—Sea $(N^2)'$ el conjunto de pares (a, b) tales que $a \geqslant b$. La correspondencia $(a, b) \longrightarrow a - b$ define una aplicación de $(N^2)'$ en N. Tal aplicación es, evidentemente, suprayectiva, pero no inyectiva. Dos pares (a, b) y (a', b') se aplican sobre el mismo entero si

$$x = a - b = a' - b',$$

o también $a' = a + p$, $b' = b + p$, donde p designa un entero determinado. Se trata, pues, de una relación de equivalencia $\mathcal{E}'$ entre pares (véase página 16). Una clase, la correspondiente a $x \in N$, comprende el conjunto de pares de la forma $(m + x, m)$, siendo m un número cualquiera de N.

En estas condiciones, el conjunto cociente $(N^2)'/\mathcal{E}'$ se halla en correspondencia biunívoca con N (véase pág. 16). Las operaciones definidas sobre N inducen operaciones sobre el conjunto precedente que le hacen isomorfo a N (véase pág. 37). Prácticamente, estas operaciones se definen sobre pares, pero son estables respecto de $\mathcal{E}'$. Escribimos

$$(a, b) + (c, d) = (a + c, b + d)$$

porque, sobre N,

$$(a - b) + (c - d) = (a + c) - (b + d);$$

la estabilidad respecto de $\mathcal{E}'$ es evidente en virtud de la última igualdad. Escribimos también

$$(a, b) \cdot (c, d) = [(ac + bd), (ad + bc)]$$

porque, sobre N,

$$(a - b)(c - d) = (ac + bd) - (ad + bc);$$

en este caso queda asimismo justificada la estabilidad por la última igualdad. Estas operaciones establecidas por isomorfismo poseen las mismas propiedades que las correspondientes operaciones de N: conmutatividad, asociatividad, distributividad. Los elementos neutros son la clase (m, m) para la adición y la clase $(m + 1, m)$ para el producto.

8. Los enteros relativos. —Generalicemos ahora los resultados anteriores al conjunto $(a, b) \in N^2$ de pares cualesquiera. Sobre este conjunto definimos primero una relación de equivalencia $\mathcal{E}$:

$$(a, b)\mathcal{E}(a', b') \qquad \text{si, y solo si,} \qquad a + b' = b + a'.$$

Existe entonces un entero p tal que

$$a' = a \pm p, \qquad b' = b \pm p;$$

encontramos el mismo formalismo que en el caso precedente. El conjunto cociente $N^2/\mathcal{E}$ se llama conjunto de los enteros relativos. Un elemento del mismo es una clase de pares equivalentes entre sí.

Precisemos en cada clase un elemento canónico que permita representarla. Puesto que

$$(a, b)\mathcal{E}(a - 1, b - 1)\mathcal{E}(a - 2, b - 2) \ldots \mathcal{E}(a - p, b - p),$$

bajo la única condición de que sean posibles las diferencias consideradas, hay lugar a considerar tres casos:

$\alpha)$ $a > b$; se llega a $(a, b)\mathcal{E}(a-b, 0)$, que es el caso precedente; decimos que se trata de un entero positivo, escrito en general $(m, 0)$ con $m \in \mathrm{N}$.

$\beta)$ $a = b$; llegamos a $(a, a)\mathcal{E}(0, 0)$, entero relativo nulo.

$\gamma)$ $a < b$; $(a, b)\mathcal{E}(0, b-a)$, se dice en este caso que se trata de un entero negativo, y se escribe $(0, m)$.

En los casos $\alpha)$ y $\gamma)$, el entero natural m se llama *valor absoluto* del número relativo. En el caso $\beta)$ este valor absoluto es igual a cero, por definición.

Operaciones.—Sobre $\mathrm{Z} = \mathrm{N}^2/\mathcal{E}$ definimos dos operaciones, la suma y el producto, por las mismas fórmulas que en el apartado 7:

$$(a, b) + (c, d) = (a + c, b + d),$$
$$(a, b) \cdot (c, d) = (ac + bd, ad + bc).$$

Puede comprobarse la estabilidad respecto de $\mathcal{E}$, pero el cálculo es inútil. Las relaciones $\mathcal{E}'$ y $\mathcal{E}$ tienen la misma expresión formal, así como las fórmulas de definición de las operaciones. Las comprobaciones necesarias se escribirían, por consiguiente, de la misma manera y, puesto que el resultado queda asegurado en el primer caso, también lo está en el segundo.

Propiedades de las operaciones.—La observación anterior es válida para las propiedades de estas operaciones: conmutatividad, asociatividad, distributividad del producto con relación a la suma. Dichas propiedades son intuitivamente evidentes para el caso $a \geqslant b$, $c \geqslant d$, pero el desarrollo formal del cálculo es independiente de esta hipótesis, y conserva su validez en el caso general.

Finalmente, la adición hace de Z un grupo. El elemento neutro es la clase (m, m). La clase de elementos equivalentes a (a, b) tiene por opuesta la clase de elementos (b, a), ya que

$$(a, b) + (b, a) = (a + b, b + a)\mathcal{E}(0, 0).$$

Tenemos, pues, una estructura de anillo conmutativo; la multiplicación posee un elemento unidad, la clase $(m+1, m)$.

Se trata, además, de un anillo de integridad, porque la condición

$$(a, b)(c, d) = (ac + bd, ad + bc)\,\mathcal{E}(0, 0)$$

implica

$$ac + bd = ad + bc, \qquad \text{de donde,} \qquad a(c-d) = b(c-d)$$

(suponiendo $c \geqslant d$). Si $c \neq d$, se tendrá, por tanto, $a = b$, de donde $(a, b) \mathcal{R} (0, 0)$. Por consiguiente, uno de los elementos (a, b), (c, d) debe ser equivalente a $(0, 0)$.

Empleo de la notación canónica.—Precisemos el cálculo sobre las clases cuando queda caracterizada cada una de ellas por su elemento canónico. Para la suma, tendremos:

$$\begin{aligned}
&(m, 0) + (n, 0) = (m + n, 0),\\
&(0, m) + (0, n) = (0, m + n),\\
&(m, 0) + (0, n) = \begin{cases} (m - n, 0), & m > n,\\ (0, n - m), & m < n,\\ (0, 0), & m = n. \end{cases}
\end{aligned}$$

Llegamos a la regla clásica: la suma de dos números positivos es un número positivo cuyo valor absoluto es la suma de los valores absolutos, ..., etc.

Finalmente, la igualdad $(m, 0) + (0, m) = (m, m) \mathcal{R} (0, 0)$ prueba que dos números del mismo valor absoluto, uno positivo y otro negativo, son opuestos.

Para el producto, se tiene, en virtud del cálculo sobre pares,

$$\begin{aligned}
&(m, 0)(n, 0) = (0, m)(0, n) = (mn, 0),\\
&(0, m)(n, 0) = (m, 0)(0, n) = (0, mn).
\end{aligned}$$

También en este caso encontramos la regla clásica: el producto de dos números de la misma categoría (positivos o negativos) es un número positivo cuyo valor absoluto es el producto de los valores absolutos

OBSERVACIÓN.—La notación anterior se sustituye generalmente por el empleo de un signo; en lugar de $(m, 0)$ escribimos $(+m)$ y, en lugar de $(0, m)$, $(-m)$, siempre con $m \in \mathbf{N}^*$. El signo $+$ que se antepone a los números positivos carece de importancia; podríamos colocar un signo cualquiera e incluso ninguno, lo que haremos más adelante. Finalmente, $-m$ se justifica para los números negativos porque son los opuestos de los positivos.

Isomorfismo de $\mathbf{Z}^+$ y de $\mathbf{N}^*$.—El estudio hecho en el apartado 7 nos asegura que el conjunto $\mathbf{Z}^+$ de los enteros positivos es isomorfo con el conjunto $\mathbf{N}^*$ (se puede comprobar también mediante la notación reducida). Este isomorfismo justifica la asimilación que hacemos de

estos dos conjuntos [1]. Identificamos, pues, $(m, 0)$ o $(+m)$ con m y $(0, 0)$ con 0. El elemento $(0, m) = -(m, 0)$ se representa por $-m$. Resulta lógica de esta manera la notación definitiva de los enteros relativos y las conocidas reglas operatorias. El siguiente teorema resume la teoría.

TEOREMA.—*El conjunto* N *puede, salvo un isomorfismo, sumergirse en un anillo en el que un elemento cualquiera, o su opuesto, pertenece a* N.

Orden sobre Z.—Sobre el anillo Z, el subconjunto Z^+ verifica las condiciones que caracterizan a un subconjunto de elementos positivos (véase pág. 79). Existe, pues, una relación de orden total compatible con la adición; escribiremos $a \leqslant b$ cuando $b - a$ sea positivo o nulo. Encontramos de nuevo de esta forma las clásicas propiedades de las desigualdades.

EJERCICIOS:

1. Fórmese el producto de tres pares (a, b), (c, d), (e, f) de N^2 y compruébese que los tres factores desempeñan un papel simétrico. Dedúzcase de lo anterior la asociatividad del producto.

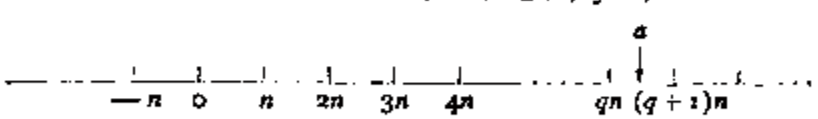
2. Compruébese directamente la ley distributiva operando sobre los pares.

3. Demuéstrese que, si aplicamos al anillo Z el método que ha permitido construirlo a partir de los enteros naturales, no obtenemos una extensión de Z, sino un nuevo anillo, isomorfo.

4. Pruébese que todo dominio de integridad que comprenda una infinidad de elementos contiene un subanillo isomorfo a Z.

EL ANILLO DE LAS CLASES RESIDUALES MODULO n

9. Congruencias.—Se puede extender el algoritmo de la división al anillo Z de los enteros relativos. Sea $a \in Z$, y n, un entero positivo.



Existen siempre dos números q y r tales que

$$a = nq + r, \qquad \text{con} \qquad 0 \leqslant r < n.$$

Sobre el gráfico anterior vemos que el entero nq tiene su imagen inmediatamente a la izquierda de la de a, o confundida con esta.

[1] El no hacerlo así nos llevaría a estudiar, con dos notaciones diferentes, dos conjuntos que gozan exactamente de las mismas propiedades.

Por definición, dos enteros a y b se llaman congruentes módulo n cuando dan el mismo resto r al ser divididos por n. Escribimos:

$$a \equiv b \text{ (mód. } n) \qquad \text{cuando} \qquad a = nq + r, \qquad b = nq' + r.$$

Ello equivale a decir que la diferencia $a - b$ es divisible por n.

TEOREMA.—*La congruencia módulo n es una relación de equivalencia sobre el anillo* Z.

Evidentemente, es reflexiva, simétrica (porque si $a - b$ es divisible por n, $b - a$ también lo es) y transitiva (si $a - b$ y $b - c$ son divisibles por n, también lo es su suma $a - c$).

TEOREMA.—*Sobre el anillo* Z *de los enteros, la adición y la multiplicación constituyen dos leyes de composición estables respecto de la congruencia módulo n.*

La demostración es inmediata:

$$a \equiv b \text{ (mód. } n) \quad \text{y} \quad c \equiv d \text{ (mód. } n) \quad \text{implican} \quad a - b \quad \text{y} \quad c - d$$

divisibles por n; por consiguiente, también lo es la suma $a + c - (b + d)$, de donde

$$a + c \equiv b + d \text{ (mód. } n).$$

Para el producto, observamos que $a = b + kn$ y $c = d + k'n$, de donde $ac = bd + k''n$; es decir, $ac \equiv bd$ (mód. n). En particular, $a \equiv b$ (mód. n) implica que cualquiera que sea c:

α) $a + c \equiv b + c$, y recíprocamente, ya que los enteros relativos forman un grupo aditivo;

β) $ac \equiv bc$, pero el recíproco no es cierto; p. ej., la congruencia $35 \equiv 15$ (mód. 10) no entraña $7 \equiv 3$ (mód. 10), porque el factor 5 por el que dividimos pertenece al módulo. Sin embargo, podemos enunciar el teorema siguiente:

TEOREMA.—*Cuando c es primo con n, la igualdad*

$$ca \equiv cb \text{ (mód. } n) \qquad \textit{implica} \qquad a \equiv b \text{ (mód. } n).$$

El número n divide a $ca - cb$ o, lo que es lo mismo, a $c(a - b)$ y, como es primo con c, divide a la diferencia $a - b$; es decir, $a \equiv b$ (mód. n).

EJERCICIOS:

1. Haciendo uso de conocidas propiedades de aritmética, resuélvanse las siguientes congruencias:

$$5x \equiv 2 \text{ (mód. 7)}, \qquad 7x \equiv 4 \text{ (mód. 10)}, \qquad 4x + 3 \equiv 1 \text{ (mód. 5)}.$$

2. Escríbase «la prueba del 9» de las operaciones, por medio de congruencias. Extráigase la raíz cuadrada, en menos de una unidad por defecto, del número 187 342 y hágase la prueba del 9 de la operación.

3. Un entero cualquiera a es tal que $a^2 \equiv 0$ ó 1 (mód. 4).

4. Demuéstrese que, si a y n son primos entre sí, la congruencia $ax \equiv b$ (módulo n) admite solución única, módulo n.

5. Para que la congruencia $ax \equiv b$ (mód. n) tenga soluciones es necesario y suficiente que el máximo común divisor r de a y n divida a b. Si se satisface esta condición, pruébese que la congruencia posee r soluciones distintas, módulo n.

6. Determínense todos los enteros x tales que $x^2 \equiv -11$ (mód. 100).

7. Resuélvase el sistema de congruencias

$$2x \equiv 3 \text{ (mód. 5)}, \qquad 5x \equiv 2 \text{ (mód. 6)}.$$

8. ¿Cuál es el resto de la división del número 17^{548} por 7?

10. Clases residuales módulo n.—De acuerdo con el estudio general hecho en el capítulo II, las clases de equivalencia asociadas a la congruencia módulo n constituyen una partición de Z: incluimos en una misma clase, llamada *clase residual*, todos los números de la forma $a+kn$ con $k \in Z$. Existen n clases, lo que equivale a decir que el conjunto cociente Z/n comprende n elementos. Cada clase puede denotarse mediante un signo, y elegimos un número particular de esta clase (generalmente, el número positivo más pequeño o nulo) colocando encima de él una barra; Z/n comprende, pues, los n elementos

$$\bar{0},\ \bar{1},\ \bar{2},\ \ldots,\ \overline{n-1}.$$

La estabilidad de la adición y de la multiplicación sobre Z respecto de la relación de equivalencia permite deducir dos leyes de composición, llamadas adición y multiplicación, entre estas n clases. Así, por ejemplo, la suma de dos clases es la clase obtenida haciendo la suma de dos enteros cualesquiera que pertenezcan a las clases consideradas, etcétera. Estas operaciones son conmutativas y asociativas, y la multiplicación es distributiva respecto a la adición [1]. Se verifican para estos n elementos todos los postulados de un anillo conmutativo; lo denotamos por Z/n. El elemento neutro para la adición, que simbolizaremos por $\bar{0}$, corresponde a los múltiplos de n. El opuesto de la clase $\bar{p}$ es la clase que denotamos por $\overline{n-p}$; si n es par, el elemento

[1] La existencia de un grupo aditivo resulta de la teoría general del capítulo anterior (véanse págs. 69 y 70), pero se demuestra fácilmente de manera directa.

designado por $(\bar{n}/2)$ es su propio opuesto. El elemento neutro para la multiplicación, denotado por $\bar{1}$, corresponde a los números de la forma $1+kn$.

EJEMPLO.—El caso más sencillo se obtiene para $n=2$. El anillo comprende dos elementos: $\bar{0}$, que corresponde a los números pares, y $\bar{1}$, que corresponde a los impares. Podemos construir las tablas de adición y de multiplicación:

adición	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

multiplicación	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$

Estas tablas expresan propiedades bien conocidas de la teoría de los enteros:

$$\text{par} + \text{par} = \text{impar} + \text{impar} = \text{par} \qquad \text{par} + \text{impar} = \text{impar}$$
$$\text{par} \times \text{par} = \text{par} \times \text{impar} = \text{par} \qquad \text{impar} \times \text{impar} = \text{impar}$$

OTRO EJEMPLO.—Estudiemos el anillo de las clases residuales módulo 6, que contiene seis elementos, $\bar{0}$, $\bar{1}$, $\bar{2}$, $\bar{3}$, $\bar{4}$, $\bar{5}$. Se obtienen las siguientes tablas:

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{5}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$

·	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{0}$	$\bar{2}$	$\bar{4}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{2}$	$\bar{0}$	$\bar{4}$	$\bar{2}$
$\bar{5}$	$\bar{0}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Divisores de cero.—$Z/6$ admite $\bar{2}$ y $\bar{3}$ como divisores de cero: 2 y 3 son los divisores estrictos de 6, y se tiene:

$$(\text{múlt. } 6+3)(\text{múlt. } 6+2) = \text{múlt. } 6.$$

Generalicemos y precisemos este resultado.

TEOREMA. *Para que Z/n sea un dominio de integridad es necesario y suficiente que n sea primo.*

Si n no es primo, admite al menos una descomposición $a \times b = n$, de donde $\bar{a} \times \bar{b} = \bar{0}$ para las clases correspondientes. Recíprocamente, si n es primo, la condición $\bar{a} \cdot \bar{b} = \bar{0}$ exige o bien $\bar{a} = \bar{0}$ o bien $\bar{b} = \bar{0}$. En efecto, en las clases de equivalencia correspondientes tendremos $a \cdot b =$ múltiplo de $n = kn$ y, por ser primo n, divide necesariamente a a o a b. De donde, o bien $\bar{a} = 0$, o bien $\bar{b} = 0$.

Los enteros relativos como dominio de operadores de un anillo.—En un anillo A denotaremos por $n \cdot x$ o nx (para $n \in Z$ y $x \in A$) la suma de n términos iguales a x si $n > 0$, el elemento nulo 0 del anillo si $n = 0$ y $-[(-n)x]$ si $n < 0$. Así, p. ej., en Z/5 tendremos

$$\bar{2} + \bar{2} + \bar{2} = \bar{3} \cdot \bar{2} = \bar{1}.$$

Si es e el elemento neutro de A se tendrá cuidado de no confundir $n \cdot e$ con n; p. ej., podemos tener $n \cdot e = 0$ con $n \neq 0$. En Z/6 se tiene $\bar{6} \cdot \bar{1} = \bar{0}$. El mínimo entero positivo *n tal que* $n \cdot e = 0$ es la *característica* del anillo A. Si no existe tal mínimo, se dice que la característica es infinita.

EJERCICIOS:

1. Fórmense las tablas de adición y de multiplicación para Z/7 y para Z/9. En el primer caso, demuéstrese que los elementos distintos de $\bar{0}$ forman un grupo multiplicativo. Lo mismo en Z/9 si exceptuamos los elementos distintos de $\bar{0}$ y del divisor de $\bar{0}$ de este anillo.

2. Hállense todas las estructuras de anillo que admiten n elementos para $n = 2, 3, 4$ y 5.

3. Estúdiese el grupo aditivo de los elementos de Z/6. ¿Es cíclico? Hállense todos los subgrupos cíclicos.

4. Determínense todos los divisores de cero de los anillos Z/12, Z/18 y Z/24.

5. Demuéstrese que, en un anillo de característica n, todo elemento a verifica a $n \cdot a = 0$, pero que (ello depende de la elección de a) n no es necesariamente el menor entero que verifica dicha condición. Cítense ejemplos.

6. Demuéstrese que la fórmula del binomio

$$(a+b)^n = a^n + C_n^1 a^{n-1} b + \ldots + C_n^n b^n,$$

donde C_n^p es el número de combinaciones de n objetos tomados p a p, es válida sobre un anillo conmutativo. En esta fórmula los enteros C_n^p desempeñan el papel de operadores. Si el anillo tiene característica n, demuéstrese que, para n primo, la fórmula se reduce a

$$(a+b)^n = a^n + b^n.$$

7. Pruébese que, en un dominio de integridad que no sea de característica 2, el único elemento igual a su opuesto es el elemento neutro para la adición.

11. Correspondencia entre Z y Z/n.—La aplicación de Z sobre Z/n definida por $x \longrightarrow \bar{x}$ es compatible con las leyes de adición y multiplicación definidas en uno y otro anillo. Los elementos que se aplican sobre la clase nula $\bar{0}$ son los múltiplos de n, y forman un subgrupo invariante N del grupo aditivo de Z, según vimos a propósito del homomorfismo de grupos. Respecto de la multiplicación, el subgrupo N posee la propiedad de «absorber» la multiplicación por un elemento *cualquiera* de Z; cuando se multiplican por un número arbitrario, los múltiplos de n siguen siendo múltiplos de n: si $x \longrightarrow \bar{0}$, $ax \longrightarrow \bar{0}$, cualquiera que sea a.

Este ejemplo da idea del homomorfismo existente entre un anillo objeto y un anillo imagen, y del papel desempeñado por el conjunto de los elementos cuya imagen es nula. En la última sección del capítulo se precisa la teoría de estos homomorfismos.

RELACION DE DIVISIBILIDAD EN LOS DOMINIOS DE INTEGRIDAD

12. La relación de divisibilidad.—Hemos visto que la relación $a \mid b$ (a divide a b), definida sobre el conjunto N de los enteros naturales, es una relación de orden.

Cuando esta relación se extiende al conjunto Z de los enteros relativos, la antisimetría desaparece; p. ej.,

$$5|(-5) \quad \text{y} \quad (-5)|5 \quad \text{con} \quad 5 \neq (-5).$$

Esto se debe a que la unidad admite el divisor (-1). Los elementos que se dividen mutuamente se asocian por pares. Una teoría general de la relación $|$ deberá examinar, por consiguiente, las clases de elementos asociados. Sobre un dominio de integridad cualquiera, definimos la relación $|$ por

$$(a|b) \iff (\exists c): ac = b.$$

TEOREMA 1.—*La relación de divisibilidad es reflexiva y transitiva.* En efecto, $a|a$, ya que $ae = a$. Por otra parte, $a|b$ y $b|c$ implican $aq = b$ y $bq' = c$; por consiguiente, $aqq' = c$.

OBSERVACIONES.—Es de señalar que $(\forall x)\ e|x$ y $x|0$. Pero 0 solo divide a él mismo, porque $0x=0(\forall x)$.

TEOREMA 2.—*Se pueden multiplicar miembro a miembro dos relaciones de divisibilidad, con la condición de no cambiar el orden de los términos.*

En efecto, $a_1|b_1$ y $a_2|b_2 \Longrightarrow a_1q_1=b_1$ y $a_2q_2=b_2$; en consecuencia,

$$a_1a_2q_1q_2=b_1b_2 \Longrightarrow a_1a_2|b_1b_2.$$

Obsérvese que es indispensable la conmutatividad.

TEOREMA 3.—*Toda combinación lineal de múltiplos de un elemento es también múltiplo de este elemento.*

Es, en efecto, evidente que $a|b_1$ y $a|b_2 \Longrightarrow a|(K_1b_1+K_2b_2)$, siendo K_1 y K_2 elementos cualesquiera del dominio. Esta propiedad se utiliza con frecuencia.

13. Elementos asociados.—Diremos *asociados* dos elementos a y b que se dividen mutuamente, relación que denotaremos por $a\mathcal{A}b$. Se trata de una relación de equivalencia; en efecto:

1) $a\mathcal{A}a$ porque $ae=a$;

2) el carácter simétrico es evidente;

3) finalmente, la relación es transitiva, ya que si $a\mathcal{A}b$ y $b\mathcal{A}c$ se tiene que $a|b$ y $b|c$, de donde $a|c$; además, $c|b$ y $b|a$, de donde $c|a$; queda así probado que $a\mathcal{A}c$.

La relación de asociación divide al dominio de integridad en clases disjuntas de elementos asociados. Examinemos, en primer lugar, la clase del elemento unidad.

Elementos unitarios.— Llamaremos elementos *unitarios* o *unidades del anillo* a los elementos asociados a la unidad y, por consiguiente, divisores de esta.

TEOREMA.—*Los elementos unitarios forman un grupo (necesariamente abeliano) para la multiplicación definida sobre el dominio.*

En efecto,

$$\epsilon_1|e \quad \text{y} \quad \epsilon_2|e \Longrightarrow \epsilon_1\epsilon_2|ee=e.$$

Por otra parte, $\epsilon|e \Longrightarrow (\exists\epsilon')$ tal que $\epsilon\epsilon'=e$. Este ϵ' es el inverso de ϵ y es unitario; evidentemente, e es el elemento neutro de este grupo.

El lector demostrará que en Z este grupo se reduce a $\{-1, +1\}$.

TEOREMA.—*Se obtienen todos los elementos de una clase de asociación multiplicando cualquiera de ellos por los elemenos unitarios.*

Sea a un elemento cualquiera y ϵ unitario; a y $a\epsilon$ son asociados, ya que $a|a\epsilon$ y $(a\epsilon)\epsilon' = a\epsilon\epsilon' = a$.

Si, por otra parte, $a \mathcal{A} b$, $a\alpha = b$ y $b\beta = a$; por consiguiente, $ba\alpha\beta = ab$, lo que demuestra que α y β son unitarios.

La partición del dominio en clases de elementos asociados precisa la relación de divisibilidad y se llega al siguiente teorema cuya demostración (muy sencilla) proponemos al lector como ejercicio.

TEOREMA.—*La relación $a|b$, definida entre los elementos, induce sobre las clases de elementos asociados una relación de divisibilidad, que es antisimétrica y determina un orden parcial.*

En la práctica se emplea un lenguaje algo diferente; se toma un representante en cada clase y se introduce la noción de divisor estricto.

DEFINICIÓN.—Un elemento a que divide al elemento b es divisor *estricto* si no es unitario, ni asociado a b.

Entonces la clase de a divide a la clase de b, de la cual difiere.

Un elemento sin divisor estricto se llama *primo*. Esta definición extiende a los dominios de integridad la noción clásica de número primo.

EJERCICIOS:

1. Demuéstrese que es indispensable la existencia del elemento unidad si deseamos que la relación | sea reflexiva. Estúdiese desde este punto de vista el anillo de los múltiplos de 2.

2. Pruébese que, sobre un anillo con divisores de 0, la relación $a \mid b$ no implica la existencia de un elemento único c tal que $ac = b$.

3. En $Z \times Z$ hacemos:

$$(a, b) + (a', b') = (a + a', b + b'),$$
$$(a, b) \cdot (a', b') = (aa' - bb', ab' + ba');$$

compruébese que se obtiene así un anillo de integridad. Demuéstrese que los elementos unitarios son $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$ y que forman un grupo cíclico de orden 4.

HOMOMORFISMO DE ANILLOS

14. Imagen homomorfa de un anillo.—Sea A un anillo, y A', un conjunto dotado de dos leyes: una adición y una multiplicación, para las cuales emplearemos las mismas notaciones corrientes en A.

Se dirá que A' es una imagen homomorfa de A por el homomorfismo h si existe una suprayección h de A sobre A' compatible con las operaciones. La imagen de una suma (o de un producto) es igual a la suma (o al producto) de las imágenes.

Consideremos el conjunto $\mathcal{J}$ de los elementos de A que tienen su imagen sobre el elemento nulo 0' de A'. La teoría de grupos nos enseña primero que $\mathcal{J}$ es, desde el punto de vista aditivo, un subgrupo distinguido del grupo aditivo de A. (En el presente caso, todo subgrupo es distinguido.) Sabemos también que A' constituye necesariamente un grupo para la adición, y que este grupo es isomorfo al grupo factor formado por las clases módulo $\mathcal{J}$.

Queda por examinar lo que aporta de nuevo la multiplicación. Se demuestra fácilmente que la multiplicación en A' es necesariamente asociativa y distributiva con relación a su adición. En otras palabras, una imagen homomorfa de un anillo es siempre un anillo. Además, si $a \in \mathcal{J}$ y $x \in A$, las imágenes de ax y de xa se forman sobre 0' porque teniendo en cuenta la compatibilidad de la aplicación y de las operaciones son, respectivamente,

$$a'x' = 0'x' = 0' \quad \text{y} \quad x'a' = x'0' = 0';$$

por consiguiente,

$$(a \in \mathcal{J}) \Longrightarrow (\forall x)\, xa \in \mathcal{J} \quad \text{y} \quad ax \in \mathcal{J}.$$

Un subgrupo aditivo de A estable para la multiplicación a la izquierda o a la derecha recibe el nombre de *ideal bilateral*.

TEOREMA.—*La imagen homomorfa de un anillo solo puede ser un anillo, y los elementos cuya imagen es nula forman un ideal bilateral.*

Antes de continuar el estudio de los homomorfismos demos un ejemplo sencillo: en Z los múltiplos de un entero n forman un ideal bilateral, porque constituyen un subgrupo aditivo, y si se multiplica un múltiplo de n por un número cualquiera, se obtiene igualmente un múltiplo de n.

Teorema de homomorfismo para los anillos.—*Cuando se aplica por ho-*

momorfismo un anillo A *sobre otro anillo* A', *el ideal bilateral* $\mathcal{J}$ *permite dividir* A *en clases mediante la relación de equivalencia* $(x-y) \in \mathcal{J}$, *y estas clases forman un anillo isomorfo a* A'.

La demostración repite palabra por palabra la construcción de Z/n, que constituye una particularización del teorema.

Desde el punto de vista aditivo,

$$x \equiv y \text{ (mód. } \mathcal{J}) \quad \text{y} \quad a \equiv b \text{ (mód. } \mathcal{J}) \Longrightarrow x+a \equiv y+b.$$

En efecto,

$$(x-y) \in \mathcal{J} \quad \text{y} \quad (a-b) \in \mathcal{J} \Longrightarrow [(x-a)-(y+b)] \in \mathcal{J},$$

obteniendo así el teorema de homomorfismo para los grupos.

La novedad interviene con la multiplicación:

$$x \equiv y \text{ (mód. } \mathcal{J}) \quad \text{y} \quad a \equiv b \text{ (mód. } \mathcal{J}) \Longrightarrow xa \equiv yb \text{ (mód. } \mathcal{J}):$$

en efecto,

$$x \equiv y \text{ (mód. } \mathcal{J}) \Longleftrightarrow (x-y) \in \mathcal{J} \Longrightarrow (\forall \lambda)\, \lambda(x-y) \in \mathcal{J}$$

en virtud de la propiedad fundamental de $\mathcal{J}$ para la multiplicación.

Por consiguiente, $\lambda x \equiv \lambda y$ (mód. $\mathcal{J}$); del mismo modo, $x\mu \equiv y\mu$.

Se tiene, pues,

$$x \equiv y \Longrightarrow xa \equiv ya,$$
$$a \equiv b \Longrightarrow ya \equiv yb,$$

de donde,

$$xa \equiv yb.$$

Las clases $\bar{x}$ módulo $\mathcal{J}$ forman, por tanto, un anillo $\bar{A}$.

$\mathcal{J} = \bar{0}$ es la clase de los elementos de imagen 0';

$\bar{x}$, clase de $(x+\mathcal{J})$, forma su imagen sobre $x'+0'=x'$;

$\bar{x} \cdot \bar{y}$, clase de $(x+\mathcal{J})(y+\mathcal{J})$, la forma sobre $(x'+0')(y'+0')=x'y'$.

Queda demostrado el isomorfismo entre $\bar{A}$ y A'. $\bar{A}$ se llama el *anillo cociente* de A por el ideal bilateral $\mathcal{J}$, y se representa por $A/\mathcal{J}$. La notación Z/n es un caso particular.

EJERCICIOS:

1. En $Z \times Z$ definimos las siguientes operaciones:

$$(a, b)+(a', b')=(a+a', b+b'),$$
$$(a, b)\cdot(a', b')=(aa', 0).$$

Demuéstrese que se tiene una estructura de anillo. Aplicamos este anillo sobre Z mediante la aplicación $(a, b) \to a$. Demuéstrese que se trata de un homomorfismo y determínese el ideal bilateral asociado. Estúdiense también los divisores de cero.

2. ¿En qué condiciones es posible aplicar Z/n sobre Z/p? Estúdiese este homomorfismo.

3. Definimos la adición en $Z \times Z$ como en el ejercicio 1, y la multiplicación, mediante $(a, b)\cdot(a', b') = (aa', ab' + ba')$. Se obtiene así un anillo. Estúdiense los divisores de cero. Estúdiese igualmente la aplicación definida por $(a, b) \to a$. ¿Es un homomorfismo? En caso afirmativo, ¿cuál es el ideal bilateral asociado?

EJERCICIOS SOBRE EL CAPITULO V

5-1. Sobre el conjunto $Z \times Z$ de pares ordenados $\alpha = (a, b)$, $\alpha' = (a', b')$ de enteros relativos se definen las siguientes leyes de composición:

$$\alpha + \alpha' = (a + a', b + b'), \qquad \alpha\alpha' = (aa' - bb', ab' + ba').$$

Demuéstrese que se obtiene de este modo un anillo. ¿Cuáles son el elemento nulo y el elemento unidad? ¿Se trata de un anillo de integridad? ¿Cuáles son los elementos unitarios?

5-2. Considérense los anillos A_1 y A_2 definidos por las siguientes leyes:

$$\begin{aligned} \alpha + \alpha' &= (a + a', b + b') && \text{para } A_1 \text{ y } A_2, \\ \alpha\alpha' &= (aa' + 2bb', ab' + ba') && \text{para } A_1, \\ \alpha\alpha' &= (aa' + 3bb', ab' + ba') && \text{para } A_2. \end{aligned}$$

Pruébese que A_1 es isomorfo al anillo del ejercicio 5-8. Demuéstrese que los anillos A_1 y A_2 no son isomorfos.

Hállense un automorfismo de A_1 y un automorfismo de A_2.

5-3. En el anillo del ejercicio 5-1 llamaremos norma de un elemento $\alpha = (a, b)$ al entero $a^2 + b^2$.

Demuéstrese que norma de $\alpha\alpha'$ = norma de $\alpha \times$ norma de α'.

Pruébese que para dos elementos α y β cualesquiera es posible definir una división con resto del tipo

$$\alpha = \chi\beta + \rho \qquad \text{norma de } \rho < \text{norma de } \beta;$$

χ es el cociente, y ρ, el resto.

5-4. Para un elemento α del anillo A_1 (ejercicio 5-2), la norma sería $a^2 - 2b^2$. ¿Se puede definir una división con resto?

5-5. Dado un grupo abeliano denotado aditivamente y cuyos elementos son a, b, etc., le dotamos de la multiplicación $ab = 0$ para todo par a, b. Demuéstrese que de esta forma se obtiene un anillo (anillo de cuadrado nulo).

En el anillo $Z/9$ el subanillo de los múltiplos de 3 es un anillo de cuadrado nulo.

5-6. Prescindimos en Z/6 de $\bar{0}$ y de los divisores de $\bar{0}$, a saber $\bar{2}$, $\bar{3}$, $\bar{4}$. Demuéstrese que las restantes clases forman un grupo para la multiplicación. ¿Cuál es su naturaleza? Hállese un grupo isomorfo ya estudiado. Considérese la misma cuestión para Z/9. Generalícese.

5-7. Se dice que un anillo A se compone de dos subanillos A' y A'' que se anulan mutuamente si:

1.° todo $a \in A$ es representable por $a=a'+a''$, con $a' \in A'$ y $a'' \in A''$;

2.° $a'a''=0$.

Demuéstrese que Z/6 es el compuesto de dos anillos que se anulan mutuamente, tales que uno es isomorfo a Z/2, y el otro, a Z/3.

5-8. 1.° Demuéstrese que el conjunto E de los números reales de la forma $a+b\sqrt{2}$, donde a y b pertenecen a Z y $\sqrt{2}$ tiene su significado habitual (número real cuyo cuadrado es igual a 2), constituyen un dominio de integridad si adoptamos como leyes de composición las reglas de adición y de multiplicación del conjunto de los números reales. ¿Cuáles son los elementos neutros para ambas operaciones?

2.° Determínense los elementos unitarios del anillo y compruébese que satisfacen a $a^2-2b^2=\pm 1$. Sea $a+b\sqrt{2}$ uno de ellos, siendo a y b positivos; demuéstrese que

$$b \leqslant a < 2b.$$

Fórmese el cociente $a_1+b_1\sqrt{2}$ del elemento considerado por el elemento unitario $1+\sqrt{2}$. Demuéstrese que si $a \neq b$, los números a_1 y b_1 verifican las desigualdades:

$$0 < a_1 < a, \qquad 0 < b_1 < b$$

y que si $a=b$, se tiene que $a_1=1$, $b_1=0$.

Concluyase de lo anterior que las sucesivas divisiones por $1+\sqrt{2}$ terminan en el número 1 después de efectuar un número finito de operaciones, y que existe un número entero positivo n tal que $a+b\sqrt{2}=(1+\sqrt{2})^n$.

5-9. En un grupo cíclico G el elemento b es de orden n. Dado el número m, primo con n, hacemos $c=b^m$. Demuéstrese que la ecuación $c^x=a$ con $a \in G$ admite siempre solución en G.

5-10. En un grupo cíclico el elemento b es de orden $n=pq$, siendo p y q primos entre sí. Demuéstrese la existencia de relaciones de la forma $b=b^{\lambda p}b^{\mu q}$, donde λ y μ designan enteros relativos.

5-11. Sea un anillo A sin elemento unidad, y Z, el anillo de los enteros relativos. El conjunto producto $Z \times A$ está formado de pares (m, a), donde $m \in Z$ y $a \in A$. Cabe definir sobre $Z \times A$ dos leyes de composición que confieran a este conjunto una estructura de anillo con elemento unidad.

a) Adición: hacemos $(m, a)+(n, b)=(m+n, a+b)$. Demuéstrese que $Z \times A$ es un grupo conmutativo con relación a esta adición.

b) Producto: definimos $(m, a) \times (n, b)=(mn, na+mb+ab)$, donde na es el n-ésimo múltiplo de a en A. Demuéstrese la asociatividad y la distributividad a la derecha y a la izquierda de la ley. Compruébese que $(1, \bar{0})$, con $\bar{0} \in A$, es el elemento neutro. Demuéstrese, finalmente, que el subconjunto A' de $Z \times A$ formado por los elementos $(0, a)$ es isomorfo a A. Dedúzcase de lo

anterior que todo anillo puede sumergirse en un anillo con elemento unidad.

5-12. *Teorema de simetrización.*—Sobre un conjunto E, defínase una ley de composición, denotada por $*$, conmutativa, asociativa y con elemento neutro. No admitiendo siempre solución sobre este conjunto la ecuación $b * x = a$, se introduce el conjunto E^2 de pares (a, b).

a) Demuéstrese que al escribir $(a, b) \equiv (a', b')$ si, y solo si, $a * b' = b * a'$ se obtiene una relación de equivalencia R sobre E^2, suponiendo los elementos regulares. Sea E^2/R el conjunto cociente.

b) Demuéstrese que la ley de composición simbolizada por $*'$ sobre E^2

$$(a, b) *' (c, d) = (a * c, b * d),$$

es estable, conmutativa y asociativa. ¿Cuál es el elemento neutro? Demuéstrese que todo elemento admite un simétrico (inverso).

c) Pruébese que existe en E^2/R un subconjunto isomorfo con el E; dicho isomorfismo hace corresponder las leyes $*'$ y $*$.

Identificando estos dos conjuntos, demuéstrese que la ecuación $b * x = a$ admite solución y que el resultado es el elemento (a, b) de E^2/R.

CAPITULO VI

ESTRUCTURA DE CUERPO

LOS AXIOMAS DE LA ESTRUCTURA DE CUERPO

1. Definición.—Se llama cuerpo todo anillo para el que los elementos no nulos forman un grupo respecto de la multiplicación. Si el grupo es abeliano, único caso que estudiaremos, el cuerpo se denomina conmutativo [1]. Los elementos x, y, z, ... de tal cuerpo, K, verifican, pues, los axiomas que siguen.

I. Grupo abeliano aditivo.

1.° Existe una ley denotada por $+$ $\quad (\forall x)(\forall y)$: $\quad (x+y) \in \mathrm{K}$.

2.° $\quad (\forall x)(\forall y) \quad : \quad x+y=y+x.$

3.° $\quad (\forall x)(\forall y)(\forall z) \quad : \quad (x+y)+z=x+(y+z).$

4.° $\quad (\exists 0)(0 \in \mathrm{K})(\forall x) \quad : \quad 0+x=x+0=x.$

5.° $\quad (\forall x)(\exists -x) \quad : \quad (-x)+x=x+(-x)=0.$

II. Grupo abeliano multiplicativo. 1.° Existe una ley que denotamos sin signo (o con $\times$, o $\cdot$) $\quad (\forall x)(\forall y)$: $\quad xy \in \mathrm{K}$.

2.° $\quad (\forall x)(\forall y) \quad : \quad xy=yx.$

3.° $\quad (\forall x)\ (\forall y)(\forall z) \quad : \quad x(yz)=(xy)z.$

4.° $\quad (\exists e)(e \in \mathrm{K})(\forall x) \quad : \quad ex=xe=x.$

5.° $\quad (\forall x,\ x \neq 0)(\exists x^{-1}) \quad : \quad x^{-1}x=xx^{-1}=e.$

III. Distributividad.

$$(\forall x)(\forall y)(\forall z) \quad : \quad x(y+z)=(y+z)x=xy+xz.$$

2. Reglas de cálculo.—Dado que un cuerpo es un anillo particular, se cumplirán las propiedades enunciadas en las páginas 76 y 77. Además, la existencia del grupo multiplicativo, nos dice que la ley de simplificación es válida para los elementos no nulos. El inverso x^{-1} posee las propiedades demostradas en las páginas 30 y 31.

---

[1] Veremos un ejemplo de cuerpo no conmutativo en la nota sobre cuaternios que se incluye al final del capítulo 10.

Cociente. TEOREMA.—*En un cuerpo cualquiera, la ecuación* $bx=a$ *admite solución única si es* $b\neq 0$.

Esta es una propiedad del grupo multiplicativo. La solución es $x=ab^{-1}$, que también se escribe $x=a/b$ y se dice que x es el cociente o razón de los dos elementos a y b; a es el numerador, y b, el denominador.

Propiedades.— 1.ª Para que sean iguales las dos razones a/b y c/d es necesario y suficiente que $ad=bc$.

La igualdad $a/b=c/d$ se escribe también $ab^{-1}=cd^{-1}$, de donde se deduce:

$$ad=a(b^{-1}b)d=(ab^{-1})bd=cd^{-1}bd=b(d^{-1}d)c=bc.$$

Inversamente, si es $ad=bc$, se tiene:

$$a/b=b^{-1}a=b^{-1}add^{-1}=b^{-1}bcd^{-1}=c/d.$$

En particular, se pueden multiplicar (o dividir) los dos términos de un cociente por un mismo elemento, distinto de cero, sin que cambie su valor:

$$a/b=ma/mb.$$

Es siempre posible reducir dos razones al mismo denominador. Señalemos, finalmente, que cualquier elemento puede considerarse como un cociente:

$$a=a/e=ma/me.$$

2.ª *Suma.*—Tenemos la fórmula $(a/b)\pm(c/d)=(ad\pm bc)/bd$.

Haciendo	$x=a/b$	e	$y=c/d$,
sabemos que	$bx=a$,		$dy=c$;
de donde se deduce que	$dbx=ad$,		$dby=bc$

y, haciendo la suma o la diferencia, $db(x\pm y)=ad\pm bc$,

$$x\pm y=(ad\pm bc)/bd.$$

3.ª *Producto.*—La fórmula es $(a/b)(c/d)=ac/bd$.

Con la misma notación y haciendo el producto, se obtiene

$$bdxy=ac, \quad \text{de donde} \quad xy=ac/bd.$$

En particular, $(a/b)(b/a)=(ab/ab)=e$, para a y $b\neq 0$.

4.ª *Cociente.*—Demostremos que $(a/b)/(c/d)=ad/bc$.

Si q es el cociente buscado tendremos $q(c/d)=a/b$ y, multiplicando por (d/c),

$$q(c/d)\,(d/c)=(a/b)\,(d/c), \quad \text{de donde} \quad q=ad/bc.$$

En resumen, vemos que son válidas sobre un cuerpo todas las reglas de cálculo establecidas en la aritmética elemental de los números enteros y de las fracciones (sean positivos o negativos) para las «cuatro operaciones» fundamentales: adición, sustracción, multiplicación y división.

EJERCICIOS:

1. Complétense las reglas precedentes demostrando las fórmulas:

1.° $(-a)^{-1}=-(a^{-1})$,
2.° $a\pm b/c=(ac\pm b)/c$,
3.° $a(b/c)=(ab/c)$,
4.° $(a/b)/c=a/(bc)$,
5.° $-(a/b)=-a/b$,
6.° $(-a)/(-b)=a/b$.

2. Todo dominio de integridad isomorfo a un cuerpo es a su vez un cuerpo.

3. Pruébese que sobre un cuerpo cualquiera, las relaciones

$$aa'=bb'=cc'$$

entre elementos no nulos, implican

$$(a+b')\,(b+c')\,(c+a')=(a'+b)\,(b'+c)\,(c'+a).$$

3. Ejemplos de cuerpos.—Indiquemos en primer lugar ejemplos constituidos por un número finito de elementos.

TEOREMA.—*Todo dominio de integridad que tenga un número finito de elementos es un cuerpo.*

Los elementos del anillo A pueden ordenarse en una sucesión

$$a_1, \quad a_2, \quad \ldots, \quad a_n, \qquad [1]$$

si suponemos que consta de n elementos. Para demostrar que A es un cuerpo, basta asociar a cada elemento $a\neq 0$ su inverso. Formemos todos los productos

$$aa_1, \quad aa_2, \quad \ldots, \quad aa_n, \qquad [2]$$

que constituyen n elementos del anillo, todos distintos porque la condición

$$aa_i=aa_j \quad \text{exigiría que fuese} \quad a_i=a_j,$$

ya que se trata de un dominio de integridad y es $a\neq 0$. El conjunto [2]

comprende, pues, una sola vez todos los elementos de A, en particular el elemento unidad e; de donde $aa_i = e$; a_i es el inverso de a.

En la práctica es suficiente construir la tabla de multiplicar del anillo para deducir de ella, de manera inmediata, el inverso de todo elemento. Hemos visto que los enteros módulo n, con n primo, forman dominios de integridad de n elementos, que son también cuerpos. El caso más sencillo se obtiene para $n=2$. Se tienen dos elementos, $\bar{0}$ y $\bar{1}$, y el inverso de $\bar{1}$ es el mismo $\bar{1}$.

Para $n=3$, hay tres elementos, $\bar{0}$, $\bar{1}$, $-\bar{1}$; los dos últimos son sus propios inversos.

Para $n=5$, se ve fácilmente que los inversos de

$$\bar{1} \quad \bar{2} \quad \bar{3} \quad \bar{4}$$

son

$$\bar{1} \quad \bar{3} \quad \bar{2} \quad \bar{4}$$

Aplicación.—En las páginas 87 y 88 hemos propuesto como ejercicios la resolución de congruencias. Se trataba allí de aplicar los métodos clásicos de la aritmética; sin embargo, si n es primo, la congruencia

$$ax \equiv b \qquad (\text{mód. } n)$$

se reduce a la ecuación $ax=b$ sobre el cuerpo de los enteros módulo n. Su solución es $x=a^{-1}b=b/a$. De esta forma, el lector puede hallar de nuevo los resultados de los ejercicios indicados.

El método se puede ampliar al caso de sistemas de congruencias con varias incógnitas, todas respecto al mismo módulo n. El sistema de congruencias

$$\begin{aligned} a'x + b'y &\equiv c' \qquad (\text{mód. } n), \\ ax + by &\equiv c \qquad (\text{mód. } n) \end{aligned}$$

se reduce al sistema de ecuaciones

$$ax + by = c, \qquad a'x + b'y = c',$$

que se resuelve por los métodos corrientes. Las fórmulas de resolución,

$$x = \frac{cb' - bc'}{ab' - ba'}, \qquad y = \frac{ac' - ca'}{ab' - ba'}, \qquad (ab' - ba' \neq 0),$$

son las mismas, porque solo dependen de reglas de cálculo que son válidas sobre cualquier cuerpo.

EJEMPLO.—Sobre el cuerpo de los enteros módulo 5, resuélvase el sistema

$$\begin{aligned} \bar{2}x+\bar{3}y&=\bar{2},\\ \bar{1}x+\bar{2}y&=\bar{4}. \end{aligned}$$

Sumando miembro a miembro, tenemos:

$$\bar{3}x=\bar{6}=\bar{1}, \qquad \text{de donde} \qquad x=\bar{2}.$$

Multiplicando la segunda ecuación por $\bar{3}$ (opuesto de $\bar{2}$) y sumando, se obtiene:

$$\bar{9}y=\overline{14}, \qquad \text{de donde} \qquad \bar{4}y=\bar{4}, \qquad y=\bar{1}.$$

EJERCICIOS:

1. Sobre el cuerpo de los enteros módulo 7, resuélvanse los sistemas siguientes:

$$\begin{cases} -\bar{3}x+\bar{2}y=\bar{1}, \\ \bar{1}x+\bar{3}y=-\bar{2}; \end{cases} \qquad \begin{cases} \bar{2}x+\bar{3}y=-\bar{2}, \\ -\bar{1}x-\bar{2}y-\bar{3}z=\bar{0}, \\ -\bar{3}x+\bar{1}y+\bar{2}z=\bar{1}. \end{cases}$$

2. En el mismo cuerpo, resuélvase y discútase la ecuación $x^2=a$, y después la ecuación $x^2+px+q=0$ inspirándose en el método usual.

3. Demuéstrese que entre el grupo aditivo de $Z/4$ y el multiplicativo de $Z/5$ existe un isomorfismo.

4. Característica de un cuerpo. Subcuerpo.—Recordemos que en ciertos anillos es $n\cdot e=0$ para n entero y diferente de cero (véase página 90). El menor entero positivo, n, tal que $n\cdot e=0$ se llama *característica* del anillo. Si aquel no existe, convenimos en decir que la característica es infinita. Para un cuerpo, la característica puede ser infinita, como en el caso de los racionales, o finita, como en $Z/5$. En este último caso es siempre igual a un número primo; en efecto, si fuera igual a $m=p\cdot q$, se tendría $m\cdot e=(p\cdot q)\cdot e=pe\cdot qe$, de donde $pe\cdot qe=0$, lo que es imposible, ya que un cuerpo no tiene divisores de cero.

Subcuerpo. Cuerpo primo.—Un cuerpo K_1 puede hallarse contenido por completo en un cuerpo K, siendo entonces un subcuerpo de este; veremos abundantes ejemplos en las partes II y III de este libro. El menor de los subcuerpos contiene necesariamente al elemento unidad y, por consiguiente, a todos los múltiplos $n\cdot e$. Recibe el nombre de cuerpo primo de K. Si el cuerpo K tiene característica infinita, el cuerpo primo es isomorfo al cuerpo de los racionales y, por identificación, es

posible simbolizar el elemento unidad por 1. Si el cuerpo K es de característica p, el cuerpo primo es isomorfo con Z/p.

EJERCICIOS:

1. Demuéstrense de manera precisa las dos proposiciones relativas al cuerpo primo.

2. Pruébese que, sobre un cuerpo, la intersección de dos subcuerpos es también un subcuerpo.

3. ¿Cuáles son los cuerpos en los que no es aplicable el método clásico de resolución de la ecuación de segundo grado «completando el cuadrado»?

4. Demuéstrese que el cuerpo primo de un cuerpo K es la intersección de todos los subcuerpos de K.

5. Enúnciense y demuéstrense las condiciones necesarias y suficientes para que un subconjunto de un cuerpo constituya un subcuerpo.

EL CUERPO Q DE LOS NUMEROS RACIONALES

5. Introducción.—Vamos a construir ahora un cuerpo bien conocido del lector: el de los números racionales. Para evitar cualquier confusión en cuanto a las notaciones, designemos con letras griegas α, β, γ, ... los elementos del anillo de los enteros relativos, y por letras latinas a, b, c, ... los elementos de un cuerpo. El razonamiento que vamos a hacer es, además, válido para cualquier dominio de integridad conmutativo y proporciona en tal caso el *cuerpo de las fracciones* de este anillo. El método seguido es idéntico al utilizado para la construcción de los enteros relativos y podría generalizarse a otros casos. Sobre el anillo Z, la resolución de la ecuación

$$\beta x = \alpha \qquad (\beta \neq 0) \qquad [1]$$

no siempre es posible; es necesario que α sea múltiplo de β y en tal caso se tiene [1] $x = \alpha : \beta$.

Cuando se cumple la condición anterior cabe deducir una infinidad de ecuaciones que admiten la misma solución multiplicando (o dividiendo siempre que sea posible) los dos coeficientes de la ecuación [1] por un mismo entero γ:

$$x = \alpha : \beta = \alpha\gamma : \beta\gamma.$$

[1] Empleamos sistemáticamente la notación : para indicar la división sin resto en el anillo de los enteros, reservando la barra de fracción para el cociente (o razón) de dos elementos cualesquiera de un cuerpo.

Un mismo entero puede considerarse como cociente de una infinidad de pares de enteros relativos; dos pares (α, β) y (α', β') definen el mismo entero siempre que

$$\alpha : \beta = \alpha' : \beta'.$$

Esta condición es una relación de equivalencia en el conjunto de los pares (α, β). Hasta aquí hemos supuesto α divisible por β, pero se evita esta condición sin más que escribir la relación anterior en la forma

$$\alpha\beta' = \beta\alpha'$$

obtenida multiplicando ambos miembros por $\beta\beta'$.

6. Definición de números racionales.—Sea Z el anillo de los enteros relativos y consideremos el conjunto producto [1] $Z \times Z^*$; es decir, el conjunto de pares ordenados (α, β) de elementos de Z tales que $\beta \neq 0$, definiendo sobre este conjunto una relación de equivalencia [2] $\mathfrak{R}$:

$$(\alpha, \beta)\,\mathfrak{R}\,(\alpha', \beta') \quad \text{si, y solo si,} \quad \alpha\beta' = \beta\alpha'.$$

Se sabe que las clases de equivalencia asociadas a esta relación constituyen una partición del conjunto $Z \times Z^*$, y, por definición, los números racionales o fracciones algebraicas son los elementos del conjunto cociente $(Z \times Z^*)/\mathfrak{R}$.

Se ve que $(\alpha, \beta)\mathfrak{R}(\mu\alpha, \mu\beta)$, cualquiera que sea $\mu \in Z^*$. Recíprocamente, se obtiene la misma fracción algebraica dividiendo α y β por uno de sus divisores comunes, si existen estos. En el caso particular del anillo Z, podremos, pues, simplificar la fracción (α, β) dividiendo α y β por su m. c. d., pero siguiendo este camino es imposible obtener una notación simplificada análoga a la de los enteros naturales [3].

7. Operaciones.—Definiremos sobre el conjunto $Z \times Z^*$ dos leyes de composición internas estables respecto de $\mathfrak{R}$. Para ello, hay que tener en cuenta que un par (α, β), en el caso en que β divida a α, debe poseer las mismas propiedades que el entero $\alpha : \beta$. De acuerdo con las reglas de la aritmética, tendremos:

$$\alpha : \beta + \gamma : \delta = (\alpha\delta) : (\beta\delta) + (\beta\gamma) : (\beta\delta) = (\alpha\delta + \beta\gamma) : \beta\delta$$

[1] Con Z^* se designa el conjunto Z privado del elemento 0.

[2] La reflexividad y la simetría son evidentes; la transitividad se obtiene multiplicando miembro a miembro las igualdades $\alpha\beta' = \beta\alpha'$ y $\beta\alpha'' = \alpha\beta''$ y simplificando por $\alpha\beta$ el resultado.

[3] La divisibilidad es una relación de orden parcial, mientras que la desigualdad es una relación de orden total.

y

$$(\alpha : \beta) \times (\gamma : \delta) = (\alpha\gamma) : (\beta\delta).$$

De donde, las dos definiciones:

1) **Suma.**—Sobre $Z \times Z^*$ la suma de dos elementos (α, β) y (γ, δ) es

$$(\alpha, \beta) + (\gamma, \delta) = (\alpha\delta + \beta\gamma, \beta\delta).$$

Es fácil comprobar la estabilidad; para

$$\alpha\beta' = \beta\alpha' \qquad \text{y} \qquad \gamma\delta' = \delta\gamma' \tag{1}$$

debemos tener

$$(\alpha\delta + \beta\gamma)\beta'\delta' = \beta\delta(\alpha'\delta' + \beta'\gamma').$$

Escribiendo esta relación en la forma

$$\delta\delta'(\alpha\beta' - \beta\alpha') + \beta\beta'(\gamma\delta' - \delta\gamma') = 0$$

se ve que resulta de la hipótesis de equivalencia.

Deducimos de ello una ley de composición inducida sobre el conjunto cociente, que se enuncia de la manera clásica: para sumar dos fracciones se reducen a un común denominador y se suman los numeradores.

2) **Producto.**—Sobre $Z \times Z^*$ el producto de dos elementos es

$$(\alpha, \beta) \times (\gamma, \delta) = (\alpha\gamma, \beta\delta).$$

Las relaciones [1] demuestran inmediatamente la estabilidad, porque de ellas deducimos:

$$\alpha\gamma\beta'\delta' = \beta\delta\alpha'\gamma'.$$

De donde resulta la regla de multiplicación de fracciones.

8. El cuerpo Q de los números racionales.—Queda por comprobar que las leyes de composición verifican los axiomas de la estructura de cuerpo.

Para la suma, la conmutatividad es inmediata; el elemento neutro es (0, 1) y el opuesto de (α, β) es $(-\alpha, \beta)$, o también $(\alpha, -\beta)$, que definen el mismo elemento. Estudiaremos la asociatividad en el ejercicio 1.

Para el producto, son inmediatas la conmutatividad y la asociativi-

dad. El elemento neutro es (1, 1), y el inverso de (α, β), (β, α). Finalmente, en el ejercicio 2 se comprueba la distributividad. En resumen, hemos obtenido una estructura de cuerpo conmutativo: el cuerpo de los números racionales, Q.

OBSERVACIÓN IMPORTANTE.—Se ha construido este cuerpo interpretando y generalizando ciertas propiedades de la divisibilidad sobre Z; por tanto, debemos encontrar de nuevo los enteros entre los números racionales. En efecto, cuando $\alpha = \beta\gamma$, se obtiene la equivalencia $(\alpha, \beta) \mathcal{R} (\gamma, 1)$ y, como es fácil de demostrar, los elementos $(\gamma, 1)$ forman un subconjunto isomorfo a Z. La correspondencia biunívoca:

$$(\gamma, 1) \Longleftrightarrow \gamma,$$

conserva la suma y el producto:

$$(\gamma, 1) + (\gamma', 1) = (\gamma + \gamma', 1),$$
$$(\gamma, 1) \times (\gamma', 1) = (\gamma\gamma', 1).$$

¿Es posible identificar estos elementos con los números enteros? Debemos reconocer que entre un número entero y un par de números enteros existe una diferencia esencial. Sin embargo, los enteros naturales están caracterizados por un sistema de axiomas que los define, salvo un isomorfismo (véanse págs. 36 y 37), y lo mismo sucede con los enteros relativos. Como consecuencia, no se pierde ninguna propiedad de los enteros si los identificamos con los pares $(\gamma, 1)$. Consideremos ahora un elemento cualquiera (α, β) de Q. Se tiene:

$$(\alpha, \beta) = (\alpha, 1) \times (1, \beta) = (\alpha, 1) \times (\beta, 1)^{-1} = \frac{(\alpha, 1)}{(\beta, 1)}$$

según la notación general empleada para indicar el cociente en un cuerpo. Visto el isomorfismo precedente, escribiremos $(\alpha, \beta) = \alpha/\beta$, que es la notación clásica de las fracciones. Se obtiene así el teorema que sigue.

TEOREMA.—*El anillo* Z *de los enteros relativos puede sumergirse en un cuerpo* Q *de manera que todo elemento del cuerpo sea el cociente de dos enteros.*

La resolución de la ecuación $\beta x = \alpha$, cuando β no divide a α, hace uso de este teorema. Operamos sobre el cuerpo Q, lo que equivale a reemplazar la ecuación dada por la

$$(\beta, 1)x = (\alpha, 1),$$

cuya solución es

$$x=(\alpha, 1)\times(\beta, 1)^{-1}=(\alpha, \beta), \quad \text{o sea,} \quad \alpha/\beta.$$

EJERCICIOS:

1. Sobre $Z\times Z^*$, calcúlese la suma de tres elementos (α, β), (γ, δ) y (ϵ, η). Dedúzcase de ello la asociatividad de la operación.

2. Compruébese análogamente la distributividad; es decir,

$$(\alpha, \beta)[(\gamma, \delta)+(\epsilon, \eta)]=(\alpha, \beta)(\gamma, \delta)+(\alpha, \beta)(\epsilon, \eta).$$

3. Demuéstrese que el cuerpo de los racionales no admite otro subcuerpo distinto de él mismo.

4. Sobre Q la correspondencia biunívoca $x \Longleftrightarrow ax$, en la que $a \in Q$ y $a \neq 0$, es un automorfismo del grupo aditivo. Demuéstrese que en este grupo no existe otro automorfismo y dedúzcase de ello la no existencia de automorfismo sobre el cuerpo Q.

5. Compárese, paso por paso, la construcción de los racionales con la de los enteros relativos y con el ejercicio 5-12.

6. ¿Por qué no puede sumergirse el anillo Z/6 en un cuerpo de cocientes? ¿En qué medida es posible sumergir un anillo conmutativo con divisores de cero en un dominio más amplio? (Anillo de las fracciones.)

9. Cuerpo ordenado.—Por definición, un cuerpo K es ordenado si lo está cuando se considera como anillo; es decir, si contiene un subconjunto K^+ de elementos positivos cerrado con relación a las dos leyes de composición y tal que, para todo elemento $a \in K$, se tenga $a \in K^+$, o bien $-a \in K^+$, o bien $a=0$.

Demostremos que, en virtud de la construcción efectuada de los números racionales, puede ordenarse el cuerpo Q obtenido por la condición de que un elemento (α, β) sea positivo cuando lo es $\alpha\beta$; es decir, cuando ambos números α y β son positivos o los dos negativos.

En primer lugar, esta definición es estable respecto de la relación de equivalencia $\mathcal{R}$,

$$(\alpha, \beta)\,\mathcal{R}\,(\alpha', \beta') \quad \text{para} \quad \alpha\beta'=\beta\alpha'.$$

En efecto, en virtud de la regla de los signos, si α y β tienen, p. ej., el mismo signo, α' y β' serán también del mismo signo, para que puedan ser iguales los productos.

La suma de dos números positivos (α, β) y (γ, δ) es positiva. En efecto, ya que un cuadrado es siempre positivo, las desigualdades $\alpha\beta > 0$ y $\gamma\delta > 0$ implican

$$\delta^2\alpha\beta > 0 \quad \text{y} \quad \beta^2\gamma\delta > 0,$$

de donde

$$\delta^2\alpha\beta+\beta^2\gamma\delta=(\alpha\delta+\beta\gamma)\beta\delta>0.$$

Los dos términos de la suma $(\alpha\delta+\beta\gamma,\ \beta\delta)$ tienen, por consiguiente, el mismo signo. El producto de dos elementos positivos es también positivo, porque $\alpha\gamma\times\beta\delta>0$. Finalmente, un elemento cualquiera $(\alpha,\ \beta)$ será nulo si $\alpha=0$, positivo si α y β son del mismo signo, y negativo si α y β son de signos contrarios. En este último caso, el opuesto $(-\alpha,\ \beta)\equiv(\alpha,\ -\beta)$ es positivo.

Se ve de este modo que el cuerpo Q de los números racionales es ordenado. Insistimos en que el orden obtenido respeta el orden establecido en el anillo de los enteros positivos, ya que a todo entero positivo α corresponde en Q la fracción $(\alpha,\ 1)$ y, evidentemente, $\alpha\times 1>0$. Recíprocamente, si se desea establecer sobre Q un orden que conserve el existente sobre el anillo, caemos en la definición precedente, ya que para $\alpha>0$ debemos tener $(\alpha,\ 1)>0$. Pero $(\alpha,\ 1)\equiv(\gamma,\ \delta)$ implica $\alpha\delta=\gamma$ y los números γ y δ son del mismo signo.

Queda demostrado, pues, el teorema siguiente:

TEOREMA.—*Los números racionales forman un cuerpo ordenado si la condición* $(\alpha,\ \beta)>0$ *significa que los enteros* α *y* β *son del mismo signo. Esta es la única manera de extender el orden de los enteros relativos al cuerpo de sus fracciones racionales.*

OBSERVACIÓN.—Hemos razonado sobre el anillo Z, pero, según hemos visto, puede definirse el cuerpo de los cocientes por un dominio de integridad cualquiera. Si este es ordenado, el método precedente permite ordenar el cuerpo de los cocientes.

10. Consecuencias.—En un cuerpo ordenado se define la desigualdad[1] $a\geqslant b$ mediante la relación $a-b\geqslant 0$, siendo también válidas aquí las propiedades de las desigualdades halladas al estudiar los anillos. Tenemos, además, las propiedades siguientes:

1.° Para $a>0$, $a^{-1}>0$, porque $a\times a^{-1}=e>0$.

2.° $a/b\leqslant c/d$ implica que $abd^2\leqslant b^2cd$. Basta multiplicar los dos miembros por $b^2d^2>0$.

3.° Para $0<a<b$, se tiene que $0<b^{-1}<a^{-1}$. Basta multiplicar por $a^{-1}b^{-1}>0$ las dos desigualdades dadas.

4.° De manera análoga $a<b<0$ implica $0>a^{-1}>b^{-1}$.

---

[1] Las letras latinas designan elementos del cuerpo. Si se trata de Q, se tiene $a=(\alpha,\ \beta)$, etc.

La noción de valor absoluto, definida en un anillo ordenado, se extiende asimismo a un cuerpo ordenado. Las propiedades se conservan (véase pág. 81).

Aplicación.—Examinemos el cuerpo de los racionales y demostremos que es *arquimediano;* es decir, que, dados dos números a y b positivos tales que $0 < b < a$, es posible siempre encontrar un entero n tal que $nb > a$. Es suficiente para ello elegir $n > a/b$, cosa posible, ya que el racional a/b está siempre comprendido entre dos enteros, o es igual a uno de ellos.

EJERCICIOS:

1. Resuélvase la inecuación $1/x < a$, en un cuerpo ordenado. Discusión.
2. Generalícense las propiedades clásicas relativas a la resolución de las inecuaciones racionales en álgebra elemental.
3. Resuélvase $x^2+x-1 > 0$ en el cuerpo Q de los racionales.
4. Demuéstrese que todo subcuerpo de un cuerpo ordenado es ordenado. ¿Cabe decir lo mismo de un *anillo* contenido en un cuerpo ordenado?
5. Pruébese que en un cuerpo ordenado, si es $a < b$, existen infinitos elementos x tales que $a < x < b$.
6. Si ϵ es un número cualquiera del cuerpo de los racionales, demuéstrese que es posible hallar un n tal que $1/10^n < \epsilon$.

11. Convergencia sobre un cuerpo ordenado.—Se dice que una sucesión indefinida de elementos pertenecientes a un cuerpo ordenado K,

$$x_1, \quad x_2, \quad \ldots, \quad x_n, \quad \ldots,$$

tiende hacia un límite a de K si, y solo si, a todo $\epsilon > 0$ de K corresponde un entero $N = N(\epsilon)$ tal que $n > N$ implica $|x_n - a| < \epsilon$.

Se dice también que la sucesión es *convergente* y se escribe $x_n \longrightarrow a$. Así, p. ej., la sucesión 0,9, 0,99, 0,999, ... tiende hacia 1 (véase ejercicio 6 anterior).

TEOREMA.—*Sobre un cuerpo ordenado, toda sucesión convergente* $x_n \longrightarrow a$ *verifica la condición de Cauchy: a todo* ϵ *positivo se puede hacer corresponder un entero* $N_1(\epsilon)$ *tal que, para* $n > N_1$ *y* $m > N_1$, *se* tenga $|x_n - x_m| < \epsilon$.

Puesto que

$$|x_n - x_m| = |(x_n - a) + (a - x_m)| < |x_n - a| + |x_m - a|,$$

basta hacer que los dos términos de la última suma sean inferiores a $\frac{\epsilon}{2}$.

Esto es posible, en virtud de la definición de límite, eligiendo n y m superiores a $N\left(\frac{\epsilon}{2}\right)$. El número $N_1(\epsilon)$ puede, pues, hacerse igual a $N\left(\frac{\epsilon}{2}\right)$.

Observemos que el recíproco de este teorema no siempre es cierto sobre un cuerpo ordenado: una sucesión indefinida que verifica la condición de Cauchy carece, en general, de límite. En particular, así sucede en el cuerpo de los racionales; basta considerar la sucesión

$$1{,}41, \quad 1{,}414, \quad 1{,}4142\ldots$$

obtenida prosiguiendo indefinidamente el algoritmo de la extracción de la raíz cuadrada de 2. Es una sucesión que verifica la condición de Cauchy y que, sin embargo, carece de límite sobre el cuerpo de los racionales. (Véanse ejercicios.)

Indiquemos finalmente que en la construcción de los números reales buscaremos un cuerpo en el que sea cierto el teorema recíproco anterior.

EJERCICIOS:

1. Indíquese por qué la sucesión precedente verifica la condición de Cauchy.

2. Demuéstrese, sobre un cuerpo ordenado, que si la sucesión $x_n \to a$, la sucesión $x_n^2 \to a^2$.

3. Demuéstrese que, si la sucesión x_n del texto tendiera hacia un límite a sobre el cuerpo de los números racionales, se tendría $a^2=2$, lo que es imposible.

4. ¿Qué podemos decir de las dos sucesiones x_n e y_n que tienden hacia el mismo número a, en el cuerpo de los racionales? Demuéstrese la recíproca de la propiedad enunciada.

EJERCICIOS SOBRE EL CAPITULO VI

6-1. Sea p un número primo, y m/n, todos los números racionales tales que n es primo con p. Demuéstrese que estos racionales forman un anillo de integridad. ¿Cuál es su cuerpo de fracciones?

6-2. Dentro del cuerpo de los racionales fórmese el menor subconjunto que sea un dominio de integridad y contenga el número 1/5. La misma cuestión si los dos números 1/5 y 1/7 deben pertenecer al subconjunto. Generalícese.

6-3. Consideremos los pares de racionales a y b, tomados en este orden, $\alpha=(a, b)$, con las leyes

$$\alpha+\alpha'=(a+a', b+b')$$
$$\alpha\alpha'=(aa'+2bb', ab'+a'b).$$

Demuéstrese que forman un cuerpo.

6-4. La misma cuestión para $\alpha\alpha' = (aa' - bb', ab' + ba')$.

6-5. Sea K un cuerpo de característica distinta de 2, y f, una aplicación de K en K tal que $f(x+y) = f(x) + f(y)$ para todo x y todo y en K, y tal que $f(x) \cdot f(e/x) = e$ para $x \neq 0$. Demuéstrese que $f(x^2) = (f(x))^2$. ¿Define esta aplicación un isomorfismo en K o sobre un subcuerpo de K?

6-6. Considérese el sistema de las tres ecuaciones

$$[\mathrm{I}) \quad \begin{cases} 3x - 2y + 1z = \alpha, \\ -2x + 1y + 2z = \beta, \\ 1x - 1y - 3z = \gamma. \end{cases}$$

sobre el cuerpo de los enteros módulo 29. Lo utilizaremos para establecer un código secreto a partir de la correspondencia

—	a	b	c	...	x	y	z	*	.
0	1	2	3	...	24	25	26	27	28

que, a los 29 signos de la primera línea hace corresponder los 29 elementos del cuerpo. Dada una palabra cualquiera, se completa con ayuda de los signos —, * ., de manera que el número de sus signos sea múltiplo de 3 y a continuación se la divide en grupos de tres signos. A los tres primeros de estos corresponden tres números de la segunda línea, que son los valores de x, y, z, de donde (en virtud de [I]) tres valores de α, β, γ (módulo 29) y tres nuevos signos que permiten codificar el grupo de tres letras considerado.

1.° Se desea codificar la palabra «algebre» escribiéndola alge*bre—

Respuesta:

oxzuuao*m.

2.° Dedúzcanse del sistema [I] las expresiones de x, y, z en función de α, β, γ, utilizándolas para descifrar la siguiente palabra: gswsqi.

Respuesta:

$$x = -5\alpha - 6\beta + 4\gamma, \qquad y = 9\alpha + 8\beta - 11\gamma,$$
$$z = 5\alpha + 5\beta - 5\gamma.$$

6-7. Decimos que un grupo abeliano G es ordenado si en él existe una relación de orden $\leqslant$ definida para todos los pares del grupo y compatible con la ley de composición del mismo; a saber:

$$a \leqslant b \quad \text{implica} \quad x * a \leqslant x * b \quad \text{para todo} \quad x \in G.$$

1.° Demuéstrese que el grupo aditivo formado por los elementos de un anillo A queda ordenado por la relación que induce en él el orden del anillo.

2.° Recíprocamente, fórmese un sistema de axiomas que caractericen un anillo ordenado, y complétese con el axioma siguiente: el grupo aditivo de los elementos del anillo es ordenado.

3.° Diremos que un grupo es ordenado arquimediano si, cualesquiera que sean los elementos a y b con $e < a < b$, existe un entero N tal que la sucesión $u_2 = a * a$, $u_3 = a * a * a$, $u_n = a$ compuesto n veces consigo mismo, sobrepase a b a partir de u_N.

4.° Demuéstrese que el grupo aditivo de los racionales y el grupo multiplicativo de los racionales positivos son arquimedianos.

5.° Se considera el conjunto de las funciones lineales $f(x) \equiv ax+b$, a y b racionales, con la ley de composición $f+f'=(a+a')x+(b+b')$. Ordenamos este conjunto, según el orden alfabético de los pares (a, b). Demuéstrese que se obtiene un grupo no arquimediano

6-8. Dados dos anillos A_1 y A_2, formamos su *compuesto directo*, A, estableciendo las siguientes reglas:

α) la intersección de A_1 y A_2 se reduce al cero de A;

β) $a_1a_2=0$ para $a_1 \in A_1$ y $a_2 \in A_2$;

γ) todo elemento a de A es de la forma $a=a_1+a_2$, con $a_1 \in A_1$, $a_2 \in A_2$.

1.° Se consideran los pares de racionales $\alpha=(a, b)$ y sobre el conjunto así definido establecemos las leyes de composición

$$\alpha+\alpha'=(a+a',\ b+b'),$$
$$\alpha\alpha'=(aa',\ bb').$$

Demuéstrese que forman un anillo, compuesto directo de dos cuerpos, Q_1 y Q_2, isomorfos a Q.

2.° Siendo ahora la segunda ley

$$\alpha\alpha'=(aa'+4bb',\ ab'+ba'),$$

pruébese que se obtiene todavía el compuesto directo de dos anillos. Caracterícense estos.

Se puede hacer $\rho=(0, 1)$ y descomponer $\rho^2=(4, 0)$ en una suma $\rho_1+\rho_2$ tal que $\rho_1 \cdot \rho_2=0$.

6-9. 1.° En el conjunto $Z \times A$ del ejercicio 5-11 (pág. 97) hemos definido una estructura de anillo. Demuéstrese que A es isomorfo a un ideal de este anillo.

2.° Estúdiense los ideales $\mathfrak{J}$ y los anillos cocientes $A/\mathfrak{J}$ en los diferentes anillos que han sido objeto de ejercicios.

6-10. Sea Q el cuerpo de los racionales (minúsculas latinas) y G el grupo de sustituciones de tres objetos (con la misma notación que en el capítulo 4, salvo para la sustitución neutra, que simbolizaremos por ϵ).

Consideramos el conjunto de combinaciones de la forma

$$X=a\epsilon+b\omega+c\theta+d\alpha+f\beta+g\gamma.$$

La suma viene dada por

$$X+X'=(a+a')\epsilon+\ldots+(g+g')\gamma;$$

el producto se obtiene multiplicando dos sumas y reemplazando productos tales como $\omega\alpha$, etc. por su valor tomado de la tabla del grupo. (Atención: ¡Respétese el orden!)

1.° Demuéstrese que de esta manera obtenemos un anillo no conmutativo: el anillo del grupo G sobre el cuerpo Q.

2.° Pruébese que este anillo es el compuesto directo de tres anillos. (Esta difícil cuestión la volveremos a considerar después de estudiar las matrices.)

6-11. En un anillo *no conmutativo* se hace

$$x * y=xy-yx.$$

1.º Demuéstrese que la ley $*$ es doblemente distributiva para la adición.
2.º Pruébese que es anticonmutativa:

$$x * y = -(y * x).$$

3.º Demuéstrese que

$$x * (y * z) + y * (z * x) + z * (x * y) = 0$$

y que

$$x * (y * z) - (x * y) * z = (z * x) * y.$$

(Identidad de Jacobi.)

6-12. Sea I un conjunto, y ϕ, X, Y, Z, ..., I, sus partes. Designamos por X' el complementario de X.

1.º Demuéstrese que queda definida una estructura de anillo en el conjunto de las partes de I haciendo:

$$XY = X \cap Y,$$
$$X + Y = (X \cap Y') \cup (Y \cap X').$$

2.º Demuéstrese que en este anillo se tiene siempre $X^2 = X$. En tales condiciones se dice que todo elemento es *idempotente* y que el anillo es *booleano*.

CAPITULO VII

ESTRUCTURA DE ESPACIO VECTORIAL

AXIOMATICA DE UN ESPACIO VECTORIAL

1. Introducción.—El conjunto de los vectores libres del espacio o del plano posee dos propiedades fundamentales:

a) Se les sabe sumar o restar, y las operaciones vectoriales poseen las mismas propiedades que la adición o la sustracción de números relativos.

b) Sabemos asociar a todo número real α (denominado *escalar*, por oposición a vector) y a todo vector $\mathbf{v}$, un vector $\alpha\mathbf{v}$ llamado producto de $\mathbf{v}$ por α. Esta operación satisface las propiedades distributiva y asociativa respecto del producto de escalares.

Conservemos sistemáticamente estas dos propiedades para caracterizar una nueva estructura que recibe el nombre de *espacio vectorial sobre un cuerpo* K.

2. Definición de espacio vectorial.—Si es K un cuerpo cualquiera, un conjunto E sobre el que se sabe definir una igualdad se denomina *espacio vectorial sobre* K, y sus elementos $\mathbf{v}$, $\mathbf{v}'$, ..., se llaman *vectores* cuando se verifican las condiciones siguientes.

Adición.—Existe sobre E una ley de composición interna, *adición*, que asocia a todo par $\mathbf{v}$, $\mathbf{v}'$ de elementos de E, distintos o no, un elemento $\mathbf{v}_1$ denominado su suma: $\mathbf{v}_1 = \mathbf{v} + \mathbf{v}'$.

E constituye un grupo abeliano con relación a esta operación; recordemos sus propiedades:

a) *conmutatividad:* $\mathbf{v} + \mathbf{v}' = \mathbf{v}' + \mathbf{v}$;

b) *asociatividad:* $(\mathbf{v} + \mathbf{v}') + \mathbf{v}'' = \mathbf{v} + (\mathbf{v}' + \mathbf{v}'')$;

c) *elemento neutro:* simbolizado por $\mathbf{0}$ y tal que $\mathbf{v} + \mathbf{0} = \mathbf{v}$ para todo $\mathbf{v} \in E$;

d) *opuesto:* denotado por $(-\mathbf{v})$ y tal que $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ para todo $\mathbf{v} \in E$.

Ley externa.—Existe sobre E una ley de composición externa de la que el dominio de operadores es el cuerpo K: a todo $\alpha \in K$ y a todo

$\mathbf{v} \in E$ sabemos asociar un vector $\mathbf{v}' = \alpha\mathbf{v}$, llamado *producto de* $\mathbf{v}$ por α. Se dice también que $\mathbf{v}'$ es el homotético de $\mathbf{v}$ de razón α. La operación verifica los siguientes axiomas:

a) *distributividad con relación a los escalares* $(\alpha+\beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$;
b) *distributividad con relación a los vectores:* $\alpha(\mathbf{v}+\mathbf{v}') = \alpha\mathbf{v} + \alpha\mathbf{v}'$;
c) *asociatividad para los escalares:* $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$;
d) *propiedad de ser neutro el elemento unidad de* K: $1\mathbf{v} = \mathbf{v}$.

En la geometría ordinaria, K es el cuerpo de los números reales, pero en otras cuestiones (geometría imaginaria, análisis) tiene interés operar sobre un cuerpo K cualquiera.

3. Propiedades elementales. Grupo aditivo.—De las propiedades de los grupos se deducen las consecuencias siguientes:

α) el elemento neutro $\mathbf{O}$ es único;
β) el opuesto de un elemento es único, $-(-\mathbf{v}) = \mathbf{v}$;
γ) la ecuación en x, $\mathbf{x} + \mathbf{v} = \mathbf{v}_1$, admite solución única, $\mathbf{x} = \mathbf{v}_1 + (-\mathbf{v})$, llamada diferencia entre $\mathbf{v}_1$ y $\mathbf{v}$, y escribimos $\mathbf{x} = \mathbf{v}_1 - \mathbf{v}$;
δ) es válida la ley de simplificación:

$$\mathbf{v} + \mathbf{v}' = \mathbf{v}_1 + \mathbf{v}' \quad \text{implica} \quad \mathbf{v} = \mathbf{v}_1$$

Ley externa.—Teniendo en cuenta las propiedades de la multiplicación por un escalar, resulta que:

α) $0\mathbf{v} = \mathbf{O}$; en efecto, en virtud de la distributividad,

$$\lambda\mathbf{v} - (\lambda + 0)\mathbf{v} = \lambda\mathbf{v} + 0\mathbf{v}$$

y, restando $\lambda\mathbf{v}$ a los dos miembros, queda $\mathbf{O} = 0\mathbf{v}$;
β) asimismo, $\lambda\mathbf{O} = \mathbf{O}$ cualquiera que sea el escalar λ;
γ) recíprocamente, la relación $\lambda\mathbf{v} = \mathbf{O}$ implica, bien $\lambda = 0$, bien $\mathbf{v} = \mathbf{O}$. Esta importante propiedad depende del hecho de que los escalares pertenecen a un cuerpo; en efecto, si λ no es nulo admite un inverso λ^{-1} y se tiene:

$$\lambda^{-1}(\lambda\mathbf{v}) = \lambda^{-1}\mathbf{O} = \mathbf{O}, \qquad \text{o} \qquad (\lambda^{-1}\lambda)\mathbf{v} = 1\mathbf{v} = \mathbf{v} = \mathbf{O};$$

δ) para un elemento λ cualquiera, $-(\lambda\mathbf{v}) = (-\lambda)\mathbf{v}$. Basta comprobar que $\lambda\mathbf{v} + (-\lambda)\mathbf{v} = \mathbf{O}$, lo que resulta fácilmente de la distributividad:

$$\lambda\mathbf{v} + (-\lambda)\mathbf{v} = [\lambda + (-\lambda)]\mathbf{v} = 0\mathbf{v} = \mathbf{O}.$$

4. Ejemplos de espacios vectoriales.—1.° Hemos citado anterior-

mente el conjunto de los vectores libres de la geometría ordinaria. Podemos también considerar el conjunto de vectores fijos del mismo origen O, como un espacio vectorial sobre R. Todo vector queda en tal caso caracterizado por un punto: su extremo.

2.° Un cuerpo K es un espacio vectorial sobre sí mismo. La ley externa se confunde con la multiplicación sobre K, y se convierte en una ley interna.

3.° *Vectores de orden n sobre* K.—Consideremos el conjunto K^n, formado por sucesiones ordenadas de n elementos cualesquiera de K:

$$\mathbf{v}=(x^1, x^2, \ldots, x^n).^1$$

Dos sucesiones son iguales si, y solo si, sus elementos del mismo orden lo son. Es fácil atribuir a este conjunto la estructura de espacio vectorial. La adición es aquella operación que, a las dos sucesiones

$$\mathbf{v}=(x^1, x^2, \ldots, x^n) \qquad \text{y} \qquad \mathbf{w}=(y^1, y^2, \ldots, y^n),$$

asocia

$$\mathbf{v}+\mathbf{w}=(x^1+y^1, x^2+y^2, \ldots, x^n+y^n).$$

Es conmutativa, asociativa y admite como elemento neutro $\mathbf{O}=(0, 0, \ldots, 0)$. La multiplicación por un escalar λ es tal que

$$\lambda\mathbf{v}=(\lambda x^1, \lambda x^2, \ldots, \lambda x^n),$$

siendo inmediato comprobar que se cumplen todos los axiomas de la estructura, ya que quedan reducidos a las propiedades elementales entre los elementos de K. Este es el espacio de los vectores de orden n sobre K, representado con la notación K^n. En álgebra lineal, esta notación designa, pues, el conjunto K^n dotado de su estructura de espacio vectorial.

4.° Veremos que el conjunto de los polinomios de una variable sobre un cuerpo es un espacio vectorial.

5.° El conjunto ($\mathfrak{B}$) de las funciones reales de la variable real x, definidas y acotadas sobre un segmento (a, b), y el conjunto de las funciones definidas y continuas sobre el mismo segmento son espacios vectoriales sobre el cuerpo de los números reales. En estos ejemplos el elemento neutro es la función nula sobre el segmento. El carácter acotado o continuo se conserva cuando se efectúa la suma de dos funciones, o el producto de una de ellas por un número real.

---

[1] En la notación $(x^1, x^2, \ldots, x^n)$, los números 1, 2, ..., n son *supraíndices* y no exponentes. Precisaremos estas notaciones más adelante.

6.° Se demuestra en análisis que el conjunto de las funciones analíticas que verifican una ecuación diferencial lineal,

$$y'' + a(x)y' + b(x)y = 0,$$

constituye un espacio vectorial sobre el cuerpo de los números reales.

7.° El cuerpo de los números complejos es un espacio vectorial sobre el cuerpo de los números reales, de igual modo que todo supercuerpo de un cuerpo K es un espacio vectorial sobre K.

8.° El cuerpo de los números reales es un espacio vectorial sobre el cuerpo Q de los números racionales; $\sqrt{2}$, π, $5\sqrt{2} + \pi/3$, etc., son vectores de este espacio.

5. Producto de espacios vectoriales.—El procedimiento que nos ha permitido definir K^n a partir de K puede generalizarse para construir nuevos espacios a partir de espacios conocidos. Sean, p. ej., E y F dos espacios sobre el mismo cuerpo K, y consideremos el conjunto producto $E \times F$, formado por los pares (v, w), donde $v \in E$ y $w \in F$. Se define una adición sobre este conjunto haciendo

$$(v, w) + (v', w') = (v + v', w + w').$$

De esta manera el conjunto $E \times F$ pasa a ser un grupo aditivo, como es fácil de comprobar, con $(\mathbf{O}_E, \mathbf{O}_F)$ como elemento neutro. Finalmente, la ley externa definida por

$$\lambda(v, w) = (\lambda v, \lambda w)$$

verifica las propiedades $\alpha)$, $\beta)$, $\gamma)$, $\delta)$, indicadas en la página 116. El conjunto $E \times F$ es, por consiguiente, un espacio vectorial. Se define de manera análoga el producto de un número finito arbitrario de espacios vectoriales.

6. Combinaciones lineales.—Un número cualquiera de vectores, $v_1, v_2, \ldots, v_n$, constituye una *familia finita*, o un *sistema* S, denotado por $S = \{v_1, v_2, \ldots, v_n\}$. Tal sistema es ordenado, pero muchas de sus propiedades son independientes del orden. Sea una familia $\{v_1, v_2, \ldots, v_n\}$, y consideremos el vector

$$v = \lambda^1 v_1 + \lambda^2 v_2 + \ldots + \lambda^n v_n = \sum_1^n \lambda^\alpha v_\alpha,$$

donde los escalares $\lambda^1, \lambda^2, \ldots, \lambda^n$, denominados coeficientes, son ele-

mentos cualesquiera de K. Decimos que $\mathbf{v}$ es una *combinación lineal* de los vectores de la familia, o también que $\mathbf{v}$ depende linealmente de estos vectores. Precisemos este concepto.

1.° El vector nulo es combinación lineal de una familia cualquiera; basta elegir coeficientes nulos.

2.° Todo vector es combinación lineal de sí mismo y, en general, de toda familia que lo contiene; en efecto, $\mathbf{v} = 1\mathbf{v}$ y, si hay otros vectores, los afectamos de coeficientes nulos.

3.° Si $\mathbf{v}$ es combinación lineal de $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ y si cada $\mathbf{v}_i$ (i varía de 1 a n) es combinación lineal de $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_p\}$, el vector $\mathbf{v}$ es combinación lineal de los vectores de esta última familia; en efecto,

$$\mathbf{v}_i = \mu_i^1 \mathbf{w}_1 + \mu_i^2 \mathbf{w}_2 + \ldots + \mu_i^p \mathbf{w}_p$$

y, sustituyendo en la igualdad

$$\mathbf{v} = \sum_1^n \lambda^\alpha \mathbf{v}_\alpha,$$

se obtiene:

$$\mathbf{v} = \left(\sum_1^n \lambda^\alpha \mu_\alpha^1 \right) \mathbf{w}_1 + \left(\sum_1^n \lambda^\alpha \mu_\alpha^2 \right) \mathbf{w}_2 + \ldots + \left(\sum_1^n \lambda^\alpha \mu_\alpha^p \right) \mathbf{w}_p,$$

lo que también se escribe como una doble suma respecto de los índices repetidos α y β,

$$\mathbf{v} = \sum_{\alpha,\beta} \lambda^\alpha \mu_\alpha^\beta \mathbf{w}_\beta \qquad (\alpha \text{ de 1 a } n,\ \beta \text{ de 1 a } p).$$

OBSERVACIONES. 1.ª El lector observará que elegimos las notaciones $\lambda^1, \ldots, \lambda^n$ (supraíndices) y $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ (subíndices) de manera que en toda suma figuran repetidos un índice inferior y otro superior. Si se tiene cuidado de seguir esta regla en las fórmulas iniciales se conservará en las fórmulas finales (véase anteriormente), lo que proporciona una comprobación sistemática.

En la escritura abreviada de una suma, $\sum_1^n \lambda^\alpha \mathbf{v}_\alpha$, α se llama *índice mudo;* carece de valor determinado, pero describe un conjunto de valores enteros, y se puede reemplazar por otro signo sin cambiar el resultado. Por el contrario, en una expresión tal como $\sum_{\alpha=1}^{p} \mu_i^\alpha \mathbf{w}_\alpha$, i es un *índice libre;* para cada valor de i varía el resultado.

2.ª Cuando se conoce sin ambigüedad el conjunto de valores que describe un índice mudo, escribiremos $\sum_{\alpha}$, o también $\sum_{}$, en lugar de $\sum_{\alpha=1}^{\alpha=n}$, refiriéndose evidentemente la suma al índice repetido en la parte superior y en la inferior.

Sistemas equivalentes.—Dos sistemas de vectores se dicen equivalentes cuando toda combinación lineal de los vectores de un sistema es combinación lineal de los vectores del otro; se trata de una relación de equivalencia en el sentido estricto del término. Es fácil demostrar que dos sistemas son equivalentes si, y solo si, todo vector de cada sistema es combinación lineal de los vectores del otro.

Isomorfismo.—La noción general de isomorfismo se aplica a los espacios vectoriales. Dos espacios, E y F, son *isomorfos* si existe una biyección de E sobre F tal que:

1.° el transformado de la suma de dos vectores sea la suma de los transformados:

$$\mathbf{v} \longrightarrow \mathbf{w} \quad \text{y} \quad \mathbf{v}' \longrightarrow \mathbf{w}' \Longrightarrow \mathbf{v}+\mathbf{v}' \longrightarrow \mathbf{w}+\mathbf{w}';$$

2.° el transformado de $\lambda\mathbf{v}$ sea $\lambda\mathbf{w}$.

Para ello es necesario y suficiente que la biyección conserve las combinaciones lineales $\lambda\mathbf{v}+\mu\mathbf{v}' \longrightarrow \lambda\mathbf{w}+\mu\mathbf{w}'$.

EJERCICIOS:

1. Si es K el cuerpo de los enteros módulo 3, considérese el espacio vectorial de los vectores con n componentes sobre K.

a) ¿Cuántos vectores hay en este espacio? Siendo $\mathbf{v}$ un vector, ¿qué cabe decir del vector $\mathbf{v}+\mathbf{v}+\mathbf{v}$? ¿Puede ser un vector la suma de tres vectores iguales?

b) Hágase la tabla completa de la adición para $n=1$, $n=2$, $n=3$.

2. Constrúyase un ejemplo de vectores con dos componentes, demostrando que el axioma $1\cdot\mathbf{v}=\mathbf{v}$ no puede probarse a partir de los otros. Tómese como ejemplo un seudoproducto por un escalar, escribiendo para

$$\mathbf{v}=(x_1, x_2), \qquad \lambda * \mathbf{v}=(\lambda x_1, 0).$$

3. Constrúyase un espacio vectorial con cuatro vectores $\mathbf{O}$, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, utilizando el cuerpo de los enteros módulo 2.

4. Considérese el conjunto de las funciones definidas sobre el segmento real $0 \leqslant x \leqslant 1$ y tales que $2f(0)=f(1)$. ¿Forman un espacio vectorial sobre el cuerpo de los números reales?

5. La misma cuestión para aquellas funciones tales que $f(x)=f(1-x)$.

6. Estúdiese el mismo problema para el conjunto de los polinomios de grado inferior o igual a 4, sobre un cuerpo K (comprendido el polinomio nulo) y para todos los polinomios de grado igual a 4.

7. La misma cuestión para todas las funciones continuas y positivas.

8. Lo mismo para las funciones continuas $f(x)$ y tales que $f(1)=f(0)+1$.

9. Demuéstrese que el axioma *d)* de la ley externa puede reemplazarse por el siguiente: la condición $\alpha\mathbf{v}=\mathbf{O}$ exige sea $\alpha=0$, o bien $\mathbf{v}=\mathbf{O}$.

SUBESPACIOS VECTORIALES

7. Definición.—Una parte E′ de un espacio vectorial E sobre K constituye un *subespacio* cuando es ella misma un espacio vectorial respecto de las leyes de composición definidas sobre E. Para ello es necesario y suficiente que E′ sea cerrado respecto de las dos leyes de composición; es decir, que:

si $\mathbf{v}\in E'$ y $\mathbf{v}'\in E'$, se tenga $(\mathbf{v}+\mathbf{v}')\in E'$;
si $\mathbf{v}\in E'$ y $\alpha\in K$, sea $\alpha\mathbf{v}\in E'$.

La pertenencia a E induce entonces sobre E′ dos leyes que verifican los axiomas de la estructura.

Las dos condiciones que preceden pueden resumirse en una sola: el subconjunto E′ debe contener, al mismo tiempo que $\mathbf{v}$ y $\mathbf{v}'$, todas sus combinaciones lineales, $\lambda\mathbf{v}+\mu\mathbf{v}'$.

Señalemos que el vector nulo de E pertenece a todo subespacio E′ En efecto, $\mathbf{O}=0\mathbf{v}$, cualquiera que sea $\mathbf{v}\in E'$.

8. Ejemplos de subespacios.—1.° El vector nulo de E constituye por sí solo un subespacio vectorial. El espacio E es en sí mismo uno de sus subespacios. Estos son los subespacios impropios. Cualquier otro subespacio se llama propio.

2.° Encontraremos numerosos subespacios de K^n; p. ej., forman un subespacio los vectores cuyas p últimas coordenadas son nulas.

3.° En el conjunto de las funciones definidas sobre R, constituyen subespacios las funciones continuas, las derivables, las funciones pares, impares, etc.

4.° En la geometría ordinaria, si consideramos los vectores del mismo origen O, aquellos cuyo extremo está en un plano dado que pasa por O forman un subespacio. Igual sucede con los vectores cuyo extremo está situado sobre una recta que pasa por O.

OBSERVACIÓN.—El espacio Q^n constituye un subconjunto de R^n, pero no es un subespacio vectorial. Es cerrado respecto de la adición (sobre R^n), pero

no lo es respecto de la multiplicación por los números reales. Se dice que Q^n proviene de R^n *por restricción del cuerpo de escalares.*

9. Intersección de subespacios.—El conjunto de los vectores comunes a dos subespacios E_1 y E_2 es evidentemente cerrado con relación a las dos leyes definidas sobre E. Constituye un subespacio, llamado intersección de E_1 y E_2, y que denotaremos por $E_1 \cap E_2$. La intersección contiene al menos el elemento neutro $\mathbf{O}$ de E. Cuando dos subespacios no poseen otros vectores comunes se denominan *disjuntos* [1].

Cabe considerar también la intersección de una familia finita, o infinita, de subespacios: es el conjunto de los vectores de E que pertenecen a cada uno de estos subespacios, y también constituye un subespacio.

10. Suma de dos subespacios.—Sean E_1 y E_2 dos subespacios de E. El conjunto de los vectores w de la forma $\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2$, en donde $\mathbf{v}_1 \in E_1$ y $\mathbf{v}_2 \in E_2$, es también cerrado para las dos leyes definidas sobre E. Constituye un subespacio, que se llama *suma* de E_1 y E_2 y se simboliza por $E_1 + E_2$. No debe confundirse con la reunión de E_1 y E_2 en el sentido de la teoría de conjuntos [2]. La suma $E_1 + E_2$ contiene a E_1 y E_2 como subespacios; en efecto, $\mathbf{v}_1 = \mathbf{v}_1 + \mathbf{O}$ pertenece a $E_1 + E_2$, ya que $\mathbf{v}_1 \in E_1$ y $\mathbf{O} \in E_2$.

OBSERVACIÓN.—Un vector w de $E_1 + E_2$ puede, en general, descomponerse de varias maneras en la forma $\mathbf{v}_1 + \mathbf{v}_2$. Si

$$\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}'_1 + \mathbf{v}'_2$$

se deduce que $\mathbf{v}_1 - \mathbf{v}'_1 = \mathbf{v}'_2 - \mathbf{v}_2$, y el vector x obtenido de esta forma, que pertenece a E_1 y E_2, también pertenece a su intersección. Recíprocamente, cualquiera que sea $\mathbf{x} \in E_1 \cap E_2$, se tiene:

$$\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2 = (\mathbf{v}_1 + \mathbf{x}) + (\mathbf{v}_2 - \mathbf{x}).$$

Suma directa.—La observación anterior muestra que, si los subespacios E_1 y E_2 son disjuntos, la descomposición de un vector w de $E_1 + E_2$ según E_1 y E_2 es posible solo de una manera. Se dice entonces que la suma es *directa* y la designamos con la notación $E_1 \oplus E_2$.

---

[1] Las dos partes E_1 y E_2 no son disjuntas en el sentido del capítulo 1, ya que $\mathbf{O}$ es común; pero en este caso se trata de *subespacios* disjuntos.

[2] Esta se compone de todos los vectores que pertenecen sea a E_1 o a E_2. En general, no es un subespacio de E.

Caso de varios subespacios.—Si E_i es una familia finita de subespacios ($i = 1, 2, \ldots, p$), su suma está constituida por el subconjunto

$$w = \sum_1^p v_i \qquad \text{donde} \qquad v_i \in E_i,$$

que es un subespacio. La suma es directa si la descomposición de un vector de este subespacio según $E_1, E_2, \ldots, E_p$ es posible solo de una manera. Para ello es necesario y suficiente que la igualdad $\sum_1^p v_i = 0$ se verifique únicamente eligiendo todos los vectores v_i nulos.

11. Subespacio engendrado por un sistema de vectores.—Si es $S = \{v_1, v_2, \ldots, v_p\}$ un sistema cualquiera de E, se comprueba fácilmente que el conjunto de combinaciones lineales de los elementos de S, $v = \sum_1^p \lambda^\alpha v_\alpha$, en donde cada una de las λ^α describe el cuerpo K, es cerrado respecto de las dos leyes de composición sobre E. Constituye un subespacio, que se dice *engendrado* por el sistema S, y los vectores $v_1, v_2, \ldots, v_p$ forman un sistema de *generadores* de este subespacio.

Puede precisarse el concepto de sistemas equivalentes (véase página 120): dos sistemas son equivalentes cuando engendran el mismo subespacio.

El lector demostrará sin dificultad la siguiente proposición:

TEOREMA.—*No cambia el subespacio engendrado por un sistema dado añadiéndole nuevos generadores que sean combinaciones lineales de los primeros.*

Recíprocamente, si es S un sistema dado, cabe preguntar si será posible engendrar el mismo subespacio con un número menor de vectores. Resolveremos esta cuestión después de estudiar la independencia lineal.

Operaciones elementales.—Sea S un sistema de generadores del subespacio E'. Indiquemos ciertas transformaciones simples del sistema S que proporcionan sistemas equivalentes.

TEOREMA.—*El subespacio engendrado por un sistema no se modifica sometiendo los vectores a una o varias de las transformaciones siguientes:*

a) *cambio de orden de los generadores;*
b) *multiplicación de un generador por un escalar distinto de cero;*
c) *adición a uno de los generadores de un múltiplo de otro generador.*

Es necesario demostrar que todo vector $\mathbf{v}$, combinación del sistema inicial $S=\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m\}$, es también combinación del nuevo sistema,

$$S'=\{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_m\}.$$

Para el caso *a)* esto resulta de la conmutatividad de la adición, ya que los $\mathbf{h}$ son idénticos a los $\mathbf{g}$, salvo el orden.

Para el caso *b)* reemplazamos $\mathbf{g}_i$ por $\mathbf{h}_i=\lambda\mathbf{g}_i$ (i, entero bien determinado), sin modificar los otros generadores. Se tiene entonces:

$$\mathbf{v}=\sum_1^m\lambda^a\mathbf{g}_a=\lambda^1\mathbf{g}_1+\lambda^2\mathbf{g}_2+\ldots+\frac{\lambda^i}{\lambda}\mathbf{h}_i+\lambda^{i+1}\mathbf{g}_{i+1}+\ldots+\lambda^m\mathbf{g}_m.$$

Para *c)* reemplazamos $\mathbf{g}_i$ por $\mathbf{h}_i=\mathbf{g}_i+k\mathbf{g}_j$ ($j\neq i$) y $\mathbf{v}$ se convierte en

$$\mathbf{v}=\lambda^1\mathbf{g}_1+\ldots+\lambda^j\mathbf{h}_j+\ldots+(\lambda^i-k\lambda^j)\mathbf{g}_i+\ldots+\lambda^m\mathbf{g}_m.$$

Vemos, por consiguiente, que toda combinación lineal de S es también combinación de S'. Por otra parte, las operaciones *a), b)* y *c)* son reversibles, ya que las operaciones inversas son del mismo tipo. Los sistemas S y S' son, pues, equivalentes. Las operaciones *a), b), c)* reciben el nombre de *operaciones elementales*. La transitividad de la equivalencia $S\sim S'$ demuestra que una sucesión de operaciones elementales (tal que cada una actúa sobre el sistema de vectores que resulta de la precedente) transforma S en un sistema equivalente S_1; es decir, proporciona asimismo un sistema de generadores del subespacio considerado.

Subespacio engendrado por una parte.—Pueden generalizarse las nociones precedentes sustituyendo el sistema S por una parte cualquiera, P, de E. Se llama combinación lineal de elementos de P todo vector $\mathbf{v}$ de la forma $\mathbf{v}=\sum_1^n\lambda^a\mathbf{v}_a$, donde $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son n vectores arbitrarios de P, y n, un entero finito cualquiera. Se comprueba inmediatamente que el conjunto de las combinaciones lineales es un subespacio de E, que también en este caso se dice engendrado por P. Este subespacio está caracterizado por la propiedad siguiente:

TEOREMA.—*El subespacio engendrado por una parte* P *de* E *es la intersección de todos los subespacios de* E *que contienen a* P. *Es a la vez el menor subespacio de* E *que contiene a* P.

Sea E_P el subespacio engendrado por P. Existen subespacios E' de E que contienen a P; el mismo E es uno de ellos. Cualquier E' contiene a toda combinación lineal de un número finito de elementos de P y, en consecuencia, contiene a E_P. De ello resulta que E_P pertenece a la intersección I de todos los E'. Recíprocamente, I es un subespacio de E que contiene a P, y, por consiguiente, a E_P; es decir, $I = E_P$.

12. Subespacios suplementarios; proyecciones.—Diremos que dos subespacios E_1 y E_2 de E son suplementarios si todo $\mathbf{v} \in E$ puede escribirse de manera única en la forma

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2, \quad \text{con} \quad \mathbf{v}_1 \in E_1 \quad \text{y} \quad \mathbf{v}_2 \in E_2.$$

Esto equivale a decir que la suma de E_1 y E_2 es directa y se confunde con el espacio E entero: $E = E_1 \oplus E_2$.

El vector $\mathbf{v}_1$ (respectivamente, $\mathbf{v}_2$) se llama *proyección* de v en E_1 (respectivamente, E_2) paralelamente a E_2 (respectivamente, E_1). Se dice también que $\mathbf{v}_1$ es la componente de v en E_1, pero esta componente no queda determinada hasta después de elegir E_2.

TEOREMA.—*La proyección de un vector de* E *sobre* E_1, *paralelamente a* E_2, *define una aplicación de* E *en* E_1 *compatible con la estructura de espacio vectorial; es decir, tal que*

$$\text{pr.}\,(\lambda\mathbf{v} + \mu\mathbf{v}') = \lambda\,\text{pr.}\,\mathbf{v} + \mu\,\text{pr.}\,\mathbf{v}'.$$

De $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ y $\mathbf{v}' = \mathbf{v}'_1 + \mathbf{v}'_2$ deducimos que

$$\lambda\mathbf{v} + \mu\mathbf{v}' = (\lambda\mathbf{v}_1 + \mu\mathbf{v}'_1) + (\lambda\mathbf{v}_2 + \mu\mathbf{v}'_2)$$

y, en el segundo miembro, los vectores encerrados entre paréntesis pertenecen, respectivamente, a E_1 y E_2. Dado que la descomposición es única, la proyección de $\lambda\mathbf{v} + \mu\mathbf{v}'$ es ciertamente $\lambda\mathbf{v}_1 + \mu\mathbf{v}'_1$.

EJEMPLOS.—1.° En la geometría ordinaria, un plano y una recta (no contenida en el plano) que pase por el origen O son dos subespacios suplementarios.

2.° Sea E el espacio de las funciones $x \longrightarrow f(x)$, definidas sobre R; E_1 el subespacio de las funciones pares, $f_1(-x) = f_1(x)$, y E_2 el subespacio de las funciones impares, $f_2(-x) = -f_2(x)$. E_1 y E_2 son dos subespacios suplementarios; tienen en común la función nula, y una

función cualquiera se escribe

$$f(x)=\frac{1}{2}[f(x)+f(-x)]+\frac{1}{2}[f(x)-f(-x)],$$

es decir, como suma de una parte par y de otra impar.

Subespacios suplementarios de un subespacio dado.—Un subespacio dado E_1 admite una infinidad de subespacios suplementarios [1]. Así, p. ej., en geometría elemental, si E_1 es un plano P, se puede tomar por suplemento una recta *cualquiera* que pase por el origen O. Dos de tales rectas D y D' constituyen dos subespacios isomorfos respecto de la proyección paralela al plano [2]. Generalicemos esta propiedad.

TEOREMA.—*Si* E_2 *y* E'_2 *son dos subespacios suplementarios de un mismo subespacio* E_1, *la proyección paralelamente a* E_1 *define un isomorfismo entre* E_2 *y* E'_2.

Sea $v_2 \in E_2$. Si se proyecta sobre E'_2 paralelamente a E_1, se obtiene v'_2 tal que

$$v_2 = v_1 + v'_2, \quad \text{con} \quad v_1 \in E_1.$$

Recíprocamente, la proyección de v'_2 sobre E_2 paralelamente a E_1, da v_2, ya que $v'_2 = v_2 + (-v_1)$. Queda así establecida una aplicación biunívoca entre los dos subespacios que, por ser una proyección, conserva la estructura de espacio vectorial. Se trata, pues, de un isomorfismo.

EJERCICIOS:

1. Hállense todos los subespacios del espacio estudiado en la página 120, 1.

2. Hállense todos los subespacios de los espacios reseñados en las páginas 117 y 118, ejemplos 5 y 8.

3. En Q^3 consideramos los dos subespacios siguientes: E_1 es el conjunto de los vectores de la forma $(0, x_2, x_3)$ y E_2 admite por generadores (1, 1, 1) y (1, 2, 3). Determínese $E_1 \cap E_2$, así como $E_1 + E_2$.

4. En el espacio de los vectores de tres componentes sobre $K = Z/5$, considérese el subespacio engendrado por (2, 2, 1) y (1, 2, 0). ¿Cuántos vectores comprende? La misma cuestión para los vectores (4, 4, 2) y (1, 1, 3).

---

[1] Esta propiedad, no evidente, se demuestra en la página 138 para el caso de un espacio de n dimensiones.

[2] La correspondencia de M (sobre D) a M' (sobre D'), tal que MM' sea paralela a P es, en efecto, un isomorfismo de las estructuras vectoriales de D y D'.

Cabe también tomar para E_1 una recta Δ; E_2 y E'_2 son dos planos P_2 y P'_2. La proyección paralelamente a Δ conserva las estructuras de espacio vectorial de P_2 y P'_2.

5. En Q^3, demuéstrese que el subespacio engendrado por (1, 2, 1) y (1, 3, 2) es el mismo que el subespacio engendrado por (1, 1, 0) y (3, 8, 5).

6. Sea E_1 el subespacio engendrado por v y v' y E_2 el engendrado por v_1, v_2, v_3. Demuéstrese que E_1+E_2 está engendrado por v, v', v_1, v_2, v_3.

7. Determínense x e y para que el vector de Q^4 $(3, 2, x, y)$ pertenezca al subespacio engendrado por (1, 4, −5, 2) y (1, 2, 3, 1).

8. Se considera el subespacio E_1 de Q^3 engendrado por (1, −3, 5) y (2, 1, 4). Hállense otros dos generadores de E_1 tales que el primero tenga su primera coordenada nula, y el segundo su segunda coordenada nula. Indíquense todas las soluciones posibles.

9. ¿Podemos determinar x de manera que el vector $(-1, 2, 5, x)$ pertenezca al subespacio de Q^4 engendrado por (4, −2, 1, 7), (1, 0, 2, 4)?

10. Para que la suma de p subespacios, $E_1, E_2, \ldots, E_p$, sea directa es necesario y suficiente que, para todo $i(1, 2, \ldots, p)$,

$$E_i \cap (E_1+E_2+\ldots+E_{i-1}+E_{i+1}+\ldots+E_p)=0.$$

11. ¿Forma un subespacio el conjunto de los vectores de Q^n para los que la suma de las n coordenadas es nula? En caso afirmativo, indíquese un sistema de generadores.

DEPENDENCIA E INDEPENDENCIA LINEAL

13. Sistema ligado.—Diremos que el sistema $S=\{v_1, v_2, \ldots, v_n\}$ es ligado cuando uno al menos de los vectores dependa linealmente de los otros, es decir, sea una combinación lineal de ellos. Para que un sistema sea ligado es necesario y suficiente que exista al menos un conjunto de escalares $\lambda^1, \lambda^2, \ldots, \lambda^n$, no todos nulos, tales que

$$\lambda^1 v_1+\lambda^2 v_2+\ldots+\lambda^n v_n=\mathbf{O}. \qquad [1]$$

Esta condición es suficiente, porque, si uno de los λ es diferente de 0, el vector correspondiente depende de los otros; $\lambda^1 \neq 0$ implica

$$v_1=-\frac{1}{\lambda^1}(\lambda^2 v_2+\ldots+\lambda^n v_n). \qquad [2]$$

Es necesaria, porque [2] implica [1] con coeficientes no todos nulos (el de v_1 es 1).

Los vectores de un sistema ligado se llaman *dependientes;* [1] es una relación de dependencia. Señalemos algunas consecuencias:

1.ª Todo sistema que comprenda al vector nulo es ligado. Una

forma siempre posible de realizar [1] consiste en elegir todos los coeficientes iguales a cero, salvo el del vector nulo.

2.ª Si se añaden otros vectores a un sistema ligado, el sistema que resulta es también ligado. Queda verificada [1] si se eligen, p. ej., coeficientes nulos para los nuevos vectores.

3.ª TEOREMA.—*Para que un sistema sea ligado es necesario y suficiente que exista un subconjunto propio de dicho sistema que engendre el mismo subespacio.*

Si $\mathbf{v}_1$ es combinación lineal de los restantes vectores, $\{\mathbf{v}_2, \ldots, \mathbf{v}_n\}$ engendra el mismo subespacio que el sistema dado (véase pág. 123) y, recíprocamente, si los dos subespacios coinciden, el sistema $\{\mathbf{v}_2, \ldots, \mathbf{v}_n\}$ constituye un sistema de generadores y $\mathbf{v}_1$ depende de estos vectores.

4.ª La dependencia es una propiedad que no hace intervenir el orden de los vectores del sistema. Si tenemos esto en cuenta es posible precisar la definición.

Para que un sistema (ordenado de vectores) sea ligado es necesario y suficiente que uno de ellos sea combinación lineal de los que le preceden. Basta considerar en la relación [1] el vector correspondiente al coeficiente no nulo de mayor índice.

14. Sistema libre. La negación lógica de la noción de sistema ligado conduce al concepto de sistema *libre*. Por consiguiente, un sistema es libre cuando uno cualquiera de sus vectores no es combinación lineal de los otros, lo cual equivale a decir que la igualdad

$$\lambda^1\mathbf{v}_1 + \lambda^2\mathbf{v}_2 + \ldots + \lambda^n\mathbf{v}_n = 0 \qquad [1]$$

se verifica solo si todos los coeficientes λ^i son nulos; la única combinación lineal nula de los vectores del sistema corresponde a elegir todos los coeficientes nulos. Los vectores de un sistema libre se llaman *independientes*. Indiquemos algunas consecuencias:

1.ª Un vector $\mathbf{v}$ no nulo constituye, por sí solo, un sistema libre. En efecto, la igualdad $\lambda\mathbf{v} = \mathbf{O}$, exige o bien que sea $\mathbf{v} = \mathbf{O}$ (excluida), o bien $\lambda = 0$.

2.ª Todos los vectores de un sistema libre son distintos de $\mathbf{O}$.

3.ª Todo subconjunto elegido en un sistema libre es también un sistema libre. En efecto, si el nuevo sistema fuera ligado, añadiéndole otros vectores seguiría siendo un sistema ligado.

4.ª TEOREMA.—*En el subespacio engendrado por n generadores independientes, todo vector puede representarse de manera única como combinación lineal de los generadores.*

Si se tuviera

$$\mathbf{v} = \sum_1^n \lambda^\alpha \mathbf{v}_\alpha = \sum_1^n \lambda'^\alpha \mathbf{v}_\alpha,$$

deduciríamos que

$$\sum_1^n (\lambda^\alpha - \lambda'^\alpha)\mathbf{v}_\alpha = \mathbf{O}$$

y por ser independientes los generadores, $\mathbf{v}_i$, se tendría que $\lambda^i = \lambda'^i$: las dos descomposiciones coinciden.

5.ª TEOREMA.—*Si el sistema* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *es libre, mientras que el sistema* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{w}\}$, *obtenido añadiéndole un vector, es ligado, el vector* $\mathbf{w}$ *es combinación lineal de* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

Por ser el sistema $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{w}\}$ ligado, tendremos la igualdad

$$\lambda^1 \mathbf{v}_1 + \lambda^2 \mathbf{v}_2 + \ldots + \lambda^n \mathbf{v}_n + \mu \mathbf{w} = \mathbf{O},$$

donde $\mu \neq 0$, ya que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es libre. Será posible, pues, despejar $\mathbf{w}$, lo que demuestra el teorema.

Generalización.—Las nociones precedentes se generalizan para cualquier parte, P, de E. Tal parte se dirá *libre* si lo es toda familia finita que se pueda deducir de ella. Una parte no libre se llama ligada. Esto significa que uno al menos de sus elementos es combinación lineal de un número finito de elementos de P.

El estudio de los polinomios nos proporcionará el ejemplo de una parte libre formada por una sucesión infinita numerable de vectores: $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i, \ldots\}$.

Un criterio de independencia.—Para el espacio K^n, donde los vectores son sucesiones de n elementos de K, resulta posible dar una condición suficiente de independencia, muy utilizada en la práctica.

Sea el sistema $S = \{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m\}$, *supuesto ordenado. Supongamos que para cada valor de* k ($k = 1, 2, \ldots, m$), *el vector de número de orden* k *tiene una de sus coordenadas* (*p. ej.*, p_k) *diferente de cero, mientras que es nula la componente de orden* p_k *de todos los vectores que siguen a* $\mathbf{g}_k$. *Si los números* $p_1, p_2, \ldots, p_m$ *son distintos dos a dos, los vectores* $\mathbf{g}$ *de* S *son independientes.*

En efecto, la igualdad

$$\lambda^1 \mathbf{g}_1 + \lambda^2 \mathbf{g}_2 + \ldots + \lambda^m \mathbf{g}_m = \mathbf{O}$$

exige $\lambda^1 = 0$ para que la coordenada de orden p_1 sea nula ($\mathbf{g}_1$ es el único vector que puede tener esta componente diferente de cero). Si deseamos anular la coordenada de orden p_2 es necesario tomar $\lambda^2 = 0$, ya que esta coordenada, que podía figurar en $\mathbf{g}_1$, no interviene en razón de la elección de λ^1. Si los números $p_1, p_2, \ldots, p_m$ son distintos dos a dos, jamás encontraremos coordenadas del mismo orden y, en consecuencia, $\lambda^1 = \lambda^2 = \ldots = \lambda^m = 0$. Los vectores son independientes.

EJEMPLO.—Los cuatro vectores de Q^4:

$$\mathbf{g}_1 \ (1, 2, -1, 8), \qquad \mathbf{g}_2 = (0, 3, 5, -1),$$
$$\mathbf{g}_3 = (0, 0, -2, 6), \qquad \mathbf{g}_4 = (0, 0, 0, 1,)$$

son, ciertamente, independientes. Basta elegir $p_1 = 1$, $p_2 = 2$, $p_3 = 3$, $p_4 = 4$.

15. Teorema fundamental de la independencia.—*En el espacio E engendrado por un sistema de n generadores* $\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n\}$, *todo sistema libre* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ *admite un número p de vectores que verifica la desigualdad* $p \leqslant n$.

Vamos a formar nuevos sistemas de generadores de E en los que los vectores v reemplazarán sucesivamente a los g. El sistema $\{\mathbf{v}_1, \mathbf{g}_1, \ldots, \mathbf{g}_n\}$ engendra E, pero es evidentemente dependiente ($\mathbf{v}_1$ depende de los vectores g); en la relación así obtenida,

$$\lambda^1 \mathbf{v}_1 + \mu^1 \mathbf{g}_1 + \ldots + \mu^n \mathbf{g}_n = \mathbf{O},$$

los coeficientes μ^i no son todos nulos, pues en tal caso se tendría $\lambda^1 \neq 0$ y, como consecuencia, $\mathbf{v}_1 = \mathbf{O}$ (imposible, por pertenecer $\mathbf{v}_1$ a un sistema libre). Existe, por tanto, uno de los g que depende de los otros vectores y, como el orden es indiferente, cabe suponer que es $\mathbf{g}_1$. Si prescindimos de este generador (véase pág. 123), obtenemos el sistema $\{\mathbf{v}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n\}$ que engendra E. A continuación consideramos el sistema $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{g}_2, \ldots, \mathbf{g}_n\}$, que es dependiente. En la relación

$$\lambda'^1 \mathbf{v}_1 + \lambda'^2 \mathbf{v}_2 + \mu'^2 \mathbf{g}_2 + \ldots + \mu'^n \mathbf{g}_n = \mathbf{O}$$

no todos los coeficientes μ' son nulos, porque el sistema $\{\mathbf{v}_1, \mathbf{v}_2\}$ pertenece a un sistema libre y es él mismo libre. Uno de los g (p. ej., $\mathbf{g}_2$) será, pues, combinación de los otros vectores; lo eliminamos y tenemos de nuevo el sistema de generadores $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{g}_3, \ldots, \mathbf{g}_n\}$.

Repitiendo el proceso, los vectores g quedan sucesivamente reemplazados por los v. Si $p \leqslant n$ se obtiene finalmente el sistema de gene-

radores $\{v_1, \ldots, v_p, g_{p+1}, \ldots, g_n\}$ en donde figuran todos los vectores v. Para $p=n$, no quedan vectores g. Demostremos que es imposible la hipótesis $p > n$. El sistema $\{v_1, v_2, \ldots, v_n\}$ engendraría E y, en consecuencia, $v_{n+1}, \ldots, v_p$ serían combinaciones lineales de los vectores precedentes, lo que no es posible dada la independencia del sistema v_i. El teorema queda así demostrado.

OBSERVACIÓN.—La comparación de los dos sistemas de generadores

$$\{g_1, g_2, \ldots, g_n\}, \qquad \{v_1, v_2, \ldots, v_p, g_{p+1}, \ldots, g_n\}$$

prueba que hemos podido reemplazar p de los vectores g, convenientemente elegidos, por los p vectores v; es el teorema de la sustitución.

EJERCICIOS:

1. Los vectores $v_1, v_2, \ldots, v_n$ son independientes; ¿qué cabe decir de los $w_1=v_1$, $w_2=v_1+v_2$, $\ldots$, $w_n=v_1+v_2+\ldots+v_n$? Generalícese.

2. Demuéstrese que si v_1, v_2, v_3 son independientes, también lo son v_1+v_2, v_2+v_3, v_3+v_1.

3. Dos vectores de orden dos sobre K, (x_1, x_2) e (y_1, y_2), son independientes si, y solo si, $x_1y_2-y_1x_2 \neq 0$.

4. Dados tres vectores, v_1, v_2, v_3, diferentes de O, ¿cuáles son las condiciones necesarias y suficientes para que el subespacio engendrado por v_1 y v_2 pueda engendrarse también por v_1 y v_3?

5. Determínese una relación de dependencia a la que satisfagan los vectores

$$(-5, 2, 8, -16), \ (-5, 3, 17, -14), \ (1, 1, 11, 6) \text{ de } Q^4.$$

6. Sean S y S' dos sistemas equivalentes que contienen el mismo número de vectores. Demuéstrese que uno de ellos es libre o ligado si, y solo si, el otro verifica las mismas propiedades.

7. Sean v_1 y v_2 dos vectores de un espacio cualquiera. Demuéstrese que los vectores $w_1=a^1v_1+a^2v_2$, $w_2=b^1v_1+b^2v_2$, $w_3=c^1v_1+c^2v_2$ son dependientes. Fórmese una relación de dependencia válida cualesquiera que sean v_1 y v_2.

8. Si α, β, γ son tres constantes reales cualesquiera, demuéstrese que $\operatorname{sen}(x+\alpha)$, $\operatorname{sen}(x+\beta)$ y $\operatorname{sen}(x+\gamma)$ son tres funciones dependientes en cualquier intervalo.

9. Si $v_1, v_2, \ldots, v_r$ son r vectores independientes de Q^n, demuéstrese que son también independientes cuando se consideran como vectores de R^n.

10. Demuéstrese que los vectores 1, $\sqrt{2}$, $\sqrt{5}$, son independientes en el espacio del ejemplo 8, página 118.

ESPACIO VECTORIAL DE n DIMENSIONES

16. Base de un espacio vectorial.—Se llama base de un espacio vectorial E toda parte libre de E que engendra E. Si $B \subset E$ es una base, todo vector $\mathbf{v} \in E$ puede escribirse de manera única como combinación lineal de un número finito de elementos de B. Se demuestra (aquí lo admitiremos) que todo espacio vectorial posee al menos una base.

Encontraremos principalmente espacios que admiten bases finitas. El conjunto de los polinomios es un ejemplo de espacio vectorial con una base infinita numerable (véase pág. 214). En este último caso, podemos numerar los vectores de la base y designar esta por $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \ldots\}$. Todo vector se escribe entonces de manera única en la forma

$$\mathbf{v} = \sum_{1}^{n} x^{\alpha} e_{\alpha},$$

donde p designa un entero finito que depende del vector considerado.

Se dice que x^{α} es la *coordenada* de índice i de $\mathbf{v}$ (respecto de la base considerada), y $x^{\alpha}e_{\alpha}$ (sin suma) se llama *componente* de $\mathbf{v}$ de índice i.

17. Espacio de dimensión finita.—Un espacio vectorial E es de *dimensión finita* si posee al menos un sistema finito de generadores: $S = \{g_1, g_2, \ldots, g_p\}$. Vamos a demostrar la existencia de bases para estos espacios.

TEOREMA I.—*Sea* E *el espacio engendrado por el sistema finito*

$$S = \{g_1, \ldots, g_p\}.$$

Existe al menos una base E *extraída de* S.

Consideremos el sistema S; si es libre, en virtud de la definición anterior, constituye una base de E, y el teorema queda demostrado; si no, existe al menos un vector g_i que es combinación lineal de los otros. Suprimiéndolo, obtenemos un nuevo sistema $S' = \{g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_p\}$ que engendra también E. Apliquemos el mismo razonamiento a S'; si es libre, constituye una base; si no, suprimimos otro vector, g_j, deduciendo un tercer sistema S''. Repitiendo un número

finito de veces esta operación estamos seguros de terminar en una base de E[1] que comprende un número finito $n=p$ de vectores.

TEOREMA II.—*Sea* E *un espacio de dimensión finita. Todas las bases de* E *son finitas y comprenden un mismo número, n, de vectores.*

Sean dos bases cualesquiera B y B' de E, deducidas o no del sistema de generadores $S=\{\mathbf{g}_1, \ldots, \mathbf{g}_p\}$. El teorema fundamental de independencia prueba que cada una de ellas encierra un número finito de vectores: n para B y n' para B' (con $n \leqslant p$ y $n' \leqslant p$). Por otra parte, cada base constituye asimismo un sistema de generadores de E; consideremos B: en E hemos encontrado n' vectores independientes y, por el teorema fundamental, es $n' \leqslant n$. Cambiando el papel de las dos bases, se llega a que $n \leqslant n'$; por consiguiente, $n=n'$.

DEFINICIÓN.—El número n obtenido de esta forma se llama *dimensión* de E. Indiquemos algunas importantes consecuencias del teorema anterior:

1.ª En un espacio de n dimensiones existen sistemas libres que comprenden m vectores si $m \leqslant n$. Todo sistema de $n+1$ vectores es necesariamente ligado.

2.ª Todo sistema libre que comprende n vectores constituye una base.

3.ª Si dividimos los vectores de una base en dos subconjuntos complementarios, los subespacios engendrados por cada uno de los sistemas así definidos son suplementarios.

Rango de un sistema.—El teorema precedente se aplica en particular cuando las bases B y B' se han extraído del sistema S. Comprenderán el mismo número $n=r$ de vectores, de donde, el siguiente teorema.

TEOREMA.—*Todos los sistemas extraídos de un sistema de p vectores y que engendran el mismo subespacio comprenden un número bien determinado, r, de vectores; los* $p-r$ *vectores restantes son combinaciones lineales de los primeros.*

DEFINICIÓN.—El número r definido por el teorema anterior se llama *rango* del sistema de vectores.

18. Ejemplos.—1.° El espacio geométrico ordinario es de tres dimensiones, porque, elegidos tres vectores no coplanarios **OA**, **OB**, **OC**, sabemos que todo vector **OM** puede escribirse:

$$\mathbf{OM}=x\mathbf{OA}+y\mathbf{OB}+z\mathbf{OC}.$$

[1] Si llegáramos a suprimir todos los vectores del sistema, forzosamente serían todos nulos, y el espacio engendrado tendría dimensión nula.

2.° El espacio K^n de los vectores de orden n sobre K es de n dimensiones. En efecto, consideremos los n vectores

$$\mathbf{e}_1=(1, 0, \ldots, 0), \qquad \mathbf{e}_2=(0, 1, \ldots, 0), \ \ldots, \ \mathbf{e}_n=(0, 0, \ldots, 1)$$

en los que el i-ésimo, $\mathbf{e}_i$, tiene una componente distinta de cero e igual a 1, la de orden i. Estos vectores forman un sistema de generadores de K^n, dado que el vector $\mathbf{v}=(x^1, x^2, \ldots, x^n)$ puede escribirse:

$$\mathbf{v}=x^1\mathbf{e}_1+x^2\mathbf{e}_2+\ldots+x^n\mathbf{e}_n. \qquad [1]$$

Además, son independientes, ya que, en efecto, la igualdad

$$x^1\mathbf{e}_1+x^2\mathbf{e}_2+\ldots+x^n\mathbf{e}_n=\mathbf{O}$$

exige que sea $x^1=0$ para que la primera coordenada del primer miembro sea nula; de igual modo $x^2=\ldots=x^n=0$.

La base introducida de esta forma, asociada de manera particularmente sencilla a la definición del espacio K^n, se llama *base natural*. Su importancia se pondrá de manifiesto en la sección siguiente.

3.° Se demuestra en análisis que toda ecuación diferencial lineal y homogénea admite como soluciones un conjunto de funciones que constituyen un espacio vectorial, de dimensión igual al orden de la ecuación.

EJERCICIOS:

1. Si E y F tienen dimensiones n y p, respectivamente, demuéstrese que la dimensión del espacio producto $E \times F$ es $n+p$.

2. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ un sistema de p vectores de $Q^n \subset R^n$. Demuéstrese que el rango de este sistema es el mismo sobre Q^n y sobre R^n (véase ejercicio 9, pág. 131).

3. ¿Qué relación debe existir en R^4 entre las coordenadas de $\mathbf{v}=(x^1, x^2, x^3, x^4)$ para que este vector pertenezca al subespacio E_1 engendrado por $\mathbf{e}_1$ y $\mathbf{e}_2+\mathbf{e}_3+\mathbf{e}_4$?

Se definen tres subespacios análogos: E_2, engendrado por $\mathbf{e}_2$, $\mathbf{e}_1+\mathbf{e}_3+\mathbf{e}_4$, y E_3 y E_4, engendrados de manera análoga. Estúdiese la intersección de E_1, E_2, E_3, E_4.

4. Determínense las coordenadas de $\mathbf{v}=(1, 2, 1, 1)$ en la base

$$\mathbf{f}_1=(1, 1, 1, 1), \ \mathbf{f}_2=(1, 1, -1, -1),$$
$$\mathbf{f}_3=(1, -1, 1, -1), \ \mathbf{f}_4=(1, -1, -1, 1)$$

19. Isomorfismo. TEOREMA.—*Para que dos espacios vectoriales* E *y* F *sean isomorfos, es necesario y suficiente que tengan la misma dimensión n. En la transformación, toda base del uno se convierte en base del otro. El isomorfismo queda perfectamente determinado una vez elegidas dos bases correspondientes.*

Sean E y F de la misma dimensión n y elijamos una base $\{\mathbf{e}_i\}$ de E y otra base $\{\mathbf{f}_i\}$ (i de 1 a n) de F. La correspondencia que asocia al vector $\mathbf{v}=\Sigma x^\alpha \mathbf{e}_\alpha$ de E el $\mathbf{w}=\Sigma x^\alpha \mathbf{f}_\alpha$ de F es evidentemente una biyección de E en F. Transforma, por otra parte, $\lambda\mathbf{v}+\mu\mathbf{v}'$ en $\lambda\mathbf{w}+\mu\mathbf{w}'$; por consiguiente, es un isomorfismo.

Recíprocamente, sean E y F isomorfos y sea $\{\mathbf{e}_i\}$ (i de 1 a n) una base de E. Llamemos $\mathbf{f}_i$ el vector asociado a $\mathbf{e}_i$ en la transformación. Según la definición de isomorfismo, el vector $\mathbf{v}=\Sigma x^\alpha \mathbf{e}_\alpha$ se convierte en el $\mathbf{w}=\Sigma x^\alpha \mathbf{f}_\alpha$. Cuando $\mathbf{v}$ describe E por completo, $\mathbf{w}$ recorre por entero F, ya que la aplicación es suprayectiva; los $\mathbf{f}_i$ forman, pues, un sistema de generadores de F que es de $n' \leqslant n$ dimensiones. Pero la reciprocidad de E y F en el isomorfismo implica la posibilidad de intercambiar sus papeles; de ello se deduce que $n \leqslant n'$. Por tanto, $n=n'$ y el isomorfismo asocia los vectores de las mismas coordenadas referidos a las bases anteriores, que son correspondientes.

Caso particular.—Todo espacio F de n dimensiones sobre K es isomorfo al espacio K^n. El isomorfismo se establece una vez elegidas las bases: en F tomamos una base arbitraria $\{\mathbf{f}_i\}$, y en K^n se puede tomar la base natural. Al vector $\mathbf{V}=\Sigma x^\alpha \mathbf{f}_\alpha$ de F corresponde entonces $\mathbf{v}=(x^1, x^2, \ldots, x^n)$ de K^n.

El estudio de F, desde el punto de vista de su estructura vectorial (dependencia, independencia entre vectores, subespacios, relación de pertenencia, etc.), se identifica así con el estudio de K^n referido a su base natural. Por esta razón, esta se denomina también *base canónica* y así la designaremos generalmente.

OBSERVACIÓN.—El estudio del isomorfismo se ha hecho utilizando solo estructuras vectoriales de E y de F. Cuando se hace intervenir una definición más precisa de los vectores de dos espacios, es a veces posible, en razón de una semejanza bastante profunda, establecer entre ellos un isomorfismo *natural;* es decir, independiente de la elección de las bases. Esta circunstancia ya se ha dado (pág. 126) en el isomorfismo entre dos subespacios suplementarios de un tercero común. La veremos igualmente en el estudio del espacio dual (véase parte IV) y justificará, en este caso, una identificación de los dos espacios considerados.

20. Automorfismo. Cambio de base.—Sea E un espacio de n dimensiones, al que aplicaremos el teorema precedente suponiendo E=F, pero eligiendo bases distintas $\{\mathbf{e}_i\}$ y $\{\mathbf{f}_i\}$. La biyección que a $\mathbf{v}=\Sigma x^\alpha \mathbf{e}_\alpha$ asocia $\mathbf{v}'=\Sigma x^\alpha \mathbf{f}_\alpha$ es un *automorfismo* de la estructura vectorial de E. Esto implica que el estudio de E, realizado utilizando una base determinada, se conserve *idénticamente* cuando cambiemos de base y

reemplacemos los vectores estudiados por sus transformados en el automorfismo asociado.

EJEMPLO.—Al vector $v=(x^1, x^2, \ldots, x^n)$ de K^n asociamos el vector $w=(sx^1, sx^2, \ldots, sx^n)$, designando por s una sustitución determinada del grupo $\mathfrak{S}_n$ (sustitución sobre las coordenadas). Pruébese que v y w se corresponden en un automorfismo de K^n.

Se demostrará sin dificultad que w tiene por componentes $(x^1, x^2, \ldots, x^n)$ en la base $f_1=s^{-1}e_1, \ldots, f_n=s^{-1}e_n$, que se deduce de la base canónica $(e_1, e_2, \ldots, e_n)$ por la sustitución s^{-1}, inversa de s.

Grupo lineal.—El conjunto de los automorfismos de un espacio vectorial constituye un subgrupo del grupo de las biyecciones definidas sobre E (considerado como conjunto). Este es el subgrupo que conserva la estructura de espacio vectorial de E[1] y recibe el nombre de *grupo lineal*.

Cambio de base.—Sea $\{e_i\}$ una base de E. Toda transformación del grupo lineal queda definida por la elección de una nueva base $\{f_i\}$. Sea v un vector dado, de coordenadas $(x_1, x_2, \ldots, x_n)$ en la base inicial. Hallemos sus nuevas coordenadas en la base $\{f_i\}$.

Los nuevos vectores de la base están definidos por sus coordenadas en la antigua base,

$$f_i=\Sigma a_i^\alpha e_\alpha, \qquad (\alpha \text{ de 1 a } n) \qquad [1]$$

lo que introduce n^2 coeficientes a_i^j. Se descompone v de dos maneras:

$$v=\Sigma x^\alpha e_\alpha=\Sigma y^\beta f_\beta.$$

Reemplazamos los f_β utilizando las fórmulas [1]:

$$v=\Sigma x^\alpha e_\alpha=\Sigma\Sigma y^\beta a_\beta^\alpha e_\alpha \qquad (\text{sumas con relación a } \alpha \text{ y } \beta)$$

Identificando en ambos miembros los coeficientes de e_i en la última igualdad, resulta:

$$x^i=\Sigma a_\beta^i y^\beta \qquad (\beta \text{ de 1 a } n). \qquad [2]$$

Estas son las expresiones de las antiguas coordenadas en función de las nuevas. Si se desea obtener las fórmulas inversas, basta resolver respecto de las y^i las relaciones precedentes, las cuales forman un sistema lineal. Pero los papeles recíprocos desempeñados por las dos bases permiten dar interpretación inmediata a las fórmulas que se obtendrían. Si

$$e_i=\Sigma b_i^\alpha f_\alpha \qquad [3]$$

son las expresiones de los vectores que integran la base primitiva en

[1] Es decir, que cambia $\lambda v+\mu v_1$ en $\lambda v'+\mu v'_1$.

la nueva base, se tiene inmediatamente:

$$y^j = \Sigma b^j_\alpha x^\alpha. \qquad [4]$$

OBSERVACIÓN.—Las coordenadas a^j_i y b^j_i de los vectores de una base, con relación a la otra, están ligadas por ciertas relaciones características. En la fórmula [1] sustituyamos las expresiones [3] de los vectores $\mathbf{e}_i$; se obtiene:

$$\mathbf{f}_i = \Sigma\Sigma a^\alpha_i b^\beta_\alpha \mathbf{f}_\beta.$$

Dado que los vectores $\mathbf{f}_i$ son independientes, se deducen de ello las relaciones siguientes:

a) si $\beta = i$, el coeficiente de $\mathbf{f}_i$ en el segundo miembro debe ser igual a 1, de donde

$$\sum_\alpha a^\alpha_i b^i_\alpha = 1 \qquad (i \text{ número elegido de 1 a } n); \qquad [5]$$

b) si $\beta = j \neq i$, deberá anularse el coeficiente de $\mathbf{f}_j$, de donde

$$\sum_\alpha a^\alpha_i b^j_\alpha = 0. \qquad [5\ bis]$$

Las n^2 relaciones que así se obtienen pueden representarse por una fórmula única. Para ello se define una función (aplicación) que, a todo par (i, j) de enteros naturales inferiores o iguales a n, asocie un número determinado, 1, si $i = j$, y 0 si $i \neq j$. El valor así asociado al par (i, j) se representa por δ^j_i y recibe el nombre de *símbolo de Kronecker*. Con esta notación, las relaciones [5] y [5 bis] se escriben:

$$\Sigma a^\alpha_i b^j_\alpha = \delta^j_i.$$

Se puede, además, intercambiar los papeles de las a y de las b, de donde resultan las nuevas relaciones:

$$\Sigma a^j_\alpha b^\alpha_i = \delta^j_i.$$

21. Dimensión de un subespacio. TEOREMA.—*Si* E *es un espacio de n dimensiones, todo subespacio* E′ *de* E *posee una dimensión* $n' \leqslant n$. *La igualdad* $n' = n$ *implica que* E′ *coincide con* E.

El número de vectores de un sistema libre de E′ está acotado por n. Por tanto, todo sistema libre comprende como máximo un número determinado n' de vectores y se verifica que $n' \leqslant n$; por consiguiente, E′ es de n' dimensiones.

Si $n'=n$, los n vectores de una base de E' constituyen n vectores independientes de E, y forman, en consecuencia, una base de este espacio; E y E' coinciden.

Subespacios suplementarios.—Completemos la noción introducida en la sección 12 (pág. 125) mediante un teorema de existencia en el caso de espacios de dimensión finita.

TEOREMA.—*En un espacio de n dimensiones,* E, *todo subespacio* E' *admite al menos un subespacio suplementario. Todo subespacio suplementario tiene por dimensión $n-n'$, si es n' la dimensión de* E'.

Sea $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n'}\}$ una base de E'. Existe en E un vector $\mathbf{f}_1$ al menos tal que el sistema $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n'}, \mathbf{f}_1\}$ es libre, pues, en caso contrario, los $\mathbf{e}_i$ (i de 1 a n') formarían una base de E y se tendría $n'=n$. Se repite el razonamiento con el sistema $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n'}, \mathbf{f}_1\}$, que es libre y de orden $n'+1$. Podemos añadirle un nuevo vector $\mathbf{f}_2$ de manera que se obtenga todavía un sistema libre si $n'+1<n$. Este proceso termina cuando se hayan introducido $n-n'$ nuevos vectores, porque se habrá obtenido una base de E. Los vectores $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{n-n'}$ engendran un subespacio E'', de $n-n'$ dimensiones, que es uno de los subespacios suplementarios.

OBSERVACIÓN.—Según el teorema de la página 126, si E'' y E''' son dos subespacios suplementarios de E', existe un isomorfismo de E'' sobre E''' que hace corresponder a los vectores $\mathbf{f}_j$ (j de 1 a $n-n'$) los vectores $\mathbf{f}_j'$ tales que

$$\mathbf{f}_j'=\mathbf{f}_j+\mathbf{v}_j \qquad (j \text{ de } 1 \text{ a } n-n'),$$

donde los $\mathbf{v}_j$ son vectores arbitrarios de E'. Recíprocamente, los $\mathbf{f}_j'$ definidos mediante esta fórmula determinan un subespacio suplementario. Es posible, pues, determinar de esta manera todos los subespacios suplementarios de un subespacio dado, cuyo número es infinito.

TEOREMA DE LA BASE INCOMPLETA.—*En un espacio de n dimensiones,* E, *todo sistema de p vectores independientes puede completarse con otros $n-p$ vectores, de manera que se obtenga una base de* E.

Los p vectores engendran un subespacio E' de p dimensiones, para el que constituyen una base. Basta ahora elegir un subespacio suplementario, E'', de E' y una base de E''. El conjunto de los n vectores obtenidos forma una base de E.

22. Intersección y suma de dos subespacios.—Sean E' y E'' dos subespacios de un espacio de n dimensiones, y denotemos por $E' \cap E''$ su intersección, y por $E'+E''$ su suma. Las dimensiones de estos subespacios están ligadas por una relación independiente de n.

TEOREMA.—*Si n' y n'' son las dimensiones de los subespacios* E′ *y* E″ *e i y s, las dimensiones de* $E' \cap E''$ *y* $E' + E''$, *se tiene*:

$$n' + n'' = i + s.$$

Si es ahora $\{e_1, e_2, \ldots, e_i\}$ una base de la intersección, podemos completarla en cada uno de los subespacios E′ y E″:

$$\begin{array}{ll} \text{en } E' & \{e_1, e_2, \ldots, e_i, f_1, \ldots, f_{n'-i}\}, \\ \text{en } E'' & \{e_1, e_2, \ldots, e_i, g_1, \ldots, g_{n''-i}\}. \end{array}$$

El conjunto $\{e_1, \ldots, e_i, f_1, \ldots, f_{n'-i}, g_1, \ldots, g_{n''-i}\}$ constituye un sistema de generadores para la suma $E' + E''$. Si existiera una relación de dependencia,

$$\sum_1^i \lambda^\alpha e_\alpha + \sum_1^{n'-i} \mu^\beta f_\beta + \sum_1^{n''-i} \nu^\gamma g_\gamma = O,$$

el vector $\Sigma \nu^\gamma g_\gamma$, que por hipótesis pertenece a E″, también pertenecería a E′ y, por consiguiente, a la intersección, lo cual es imposible ya que el subespacio engendrado por los vectores g es suplementario (en E″) de la intersección. Es necesario, pues, suponer todos los coeficientes ν^γ, como igualmente los μ^β y, finalmente, los λ^α, iguales a cero. El conjunto de los vectores e, f, g constituye así una base para la suma $E' + E''$, cuyo número de dimensiones será, por consiguiente,

$$s = i + (n' - i) + (n'' - i) = n' + n'' - i.$$

Esta es la relación buscada.

Caso particular.—Cuando la suma es directa, la intersección queda reducida a $O(i = 0)$, y se tiene $s = n' + n''$. La reunión de una base de E′ y de una base de E″ da una base de $E' \oplus E''$.

ESTUDIO DE UN SUBESPACIO
SISTEMAS DE ECUACIONES LINEALES

23. Estudio de un subespacio de K^n.—Sea un espacio de n dimensiones sobre K. Salvo un isomorfismo, podemos imaginar el espacio K^n referido a su base canónica. El sistema $\{g_1, g_2, \ldots, g_m\}$, donde los vectores g están definidos por sus coordenadas, engendra un subespacio E′. Se trata de hallar una base de E′ y de precisar su dimensión. La

solución que daremos se basa en el empleo de operaciones elementales (véase pág. 124). Formamos nuevos sistemas de generadores, que se eligen de modo que satisfagan el criterio de independencia de la página 130, lo que, en general, precisa de varias etapas. Si los vectores dados son dependientes, uno de los sistemas de generadores obtenidos incluye el vector nulo, que puede desecharse sin modificar el subespacio. Supondremos los vectores con un orden determinado, pero E' es independiente de esto y el método puede aplicarse de varias maneras.

24. Determinación de una base.—Sea el sistema $\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m\}$ y designemos por $\mathbf{g}_{k_1}$ el primer generador no nulo, prescindiendo de todos los que le preceden. Una de sus componentes, la de orden p_1, p. ej., será así diferente de cero. Todos los vectores que le siguen, $\mathbf{g}_{k_1+1}, \ldots, \mathbf{g}_m$, pueden reemplazarse por combinaciones lineales

$$\mathbf{g}'_i = \mathbf{g}_i - \lambda_i \mathbf{g}_{k_1} \qquad (i > k_1)$$

eligiendo los λ_i de manera que todos los $\mathbf{g}'_i$ tengan su componente de orden p_1 igual a 0. De esta manera se forma un nuevo sistema de generadores,

$$\{\mathbf{g}_{k_1}, \mathbf{g}'_{k_1+1}, \mathbf{g}'_{k_1+2}, \ldots, \mathbf{g}'_m\},$$

que verifican parcialmente el criterio de independencia: $\mathbf{g}_{k_1}$ forma con todo vector $\mathbf{g}'$, distinto de $\mathbf{O}$, un sistema libre. Si todos los vectores $\mathbf{g}'$ son nulos, la operación ha terminado y el subespacio está engendrado por $\mathbf{g}_{k_1}$.

Si no, sea $\mathbf{g}'_{k_2}$ el primer vector $\mathbf{g}'$ no nulo, una de cuyas componentes, la de orden p_2 ($p_2 \neq p_1$), será diferente de cero.

Se forman las combinaciones

$$\mathbf{g}''_i = \mathbf{g}'_i - \mu_i \mathbf{g}'_{k_2} \qquad (i > k_2)$$

de manera que todos los vectores $\mathbf{g}''$ tengan nula la componente de orden p_2 (y también la de orden p_1). Se obtiene así un nuevo sistema,

$$\{\mathbf{g}_{k_1}, \mathbf{g}'_{k_2}, \mathbf{g}''_{k_2+1}, \ldots, \mathbf{g}''_q\} \qquad (q \leqslant m)$$

tal que $\mathbf{g}_{k_1}$, $\mathbf{g}'_{k_2}$ y uno cualquiera de los $\mathbf{g}''$ diferente de $\mathbf{O}$ sean tres vectores independientes. Si todos los $\mathbf{g}''$ son nulos el proceso queda terminado, y $\mathbf{g}_{k_1}$ y $\mathbf{g}'_{k_2}$ forman una base. En otro caso, sea $\mathbf{g}''_{k_3}$ el primer $\mathbf{g}''$ no nulo, una de cuyas componentes, la de orden p_3 ($p_3 \neq p_2$ y $p_3 \neq p_1$), es diferente de 0. Continuando la operación se obtendrá un nuevo

sistema tal que los vectores $\mathbf{g}'''$ no nulos, colocados después de $\mathbf{g}''_{k_3}$, son independientes con $\mathbf{g}_{k_1}$, $\mathbf{g}'_{k_2}$, $\mathbf{g}''_{k_3}$.

Se ve que el proceso termina cuando todos los vectores $\mathbf{g}^{(i)}$ son nulos o bien cuando se ha formado un sistema que verifica el criterio de independencia. Con m vectores $\mathbf{g}$, el número de operaciones será como máximo $m-1$. Se obtiene entonces un número r de vectores $\mathbf{g}_{k_1}$, $\mathbf{g}'_{k_2}$, $\mathbf{g}''_{k_3}$, ..., $\mathbf{g}^{(r-1)}_{k_r}$ independientes que constituyen una base de E' y este es de r dimensiones. La práctica demostrará al lector la flexibilidad del método. Las coordenadas de los vectores $\mathbf{g}$ se disponen en líneas de manera que formen una matriz [1] cuyas columnas sean las coordenadas del mismo orden. En cierto modo, la elección de los números p_1, p_2, ... es arbitraria, y casi siempre tomaremos para ellos los enteros más pequeños, pero la simplificación de ciertos cálculos puede conducir a otra elección.

Si en el sistema dado se desea precisar los vectores independientes y las expresiones de los demás, se escribe al principio de la fila la combinación efectuada en cada transformación. Cuando una fila se anula, deducimos de ello una relación de dependencia

EJEMPLO.—Determínese en Q^5 una base del subespacio engendrado por los vectores:

$$\mathbf{g}_1=(1, 2, -4, 3, 1), \qquad \mathbf{g}_2=(2, 5, -3, 4, 8),$$
$$\mathbf{g}_3=(6, 17, -7, 10, 22), \qquad \mathbf{g}_4=(1, 3, -3, 2, 0).$$

Se disponen los cálculos de la siguiente manera:

$\mathbf{g}_1$	1	2	−4	3	1		$\mathbf{g}_1$	1	2	−4	3	1
$\mathbf{g}_2$	2	5	−3	4	8		$\mathbf{g}'_2=\mathbf{g}_2-2\mathbf{g}_1$	0	1	5	−2	6
$\mathbf{g}_3$	6	17	−7	10	22		$\mathbf{g}'_3=\mathbf{g}_3-6\mathbf{g}_1$	0	5	17	−8	16
$\mathbf{g}_4$	1	3	−3	2	0		$\mathbf{g}'_4=\mathbf{g}_4-\mathbf{g}_1$	0	1	1	−1	−1
$\mathbf{g}_1$	1	2	−4	3	1		$\mathbf{g}_1$	1	2	−4	3	1
$\mathbf{g}'_2$	0	1	5	−2	6		$\mathbf{g}'_2$	0	1	5	−2	6
$\mathbf{g}''_3=\mathbf{g}'_3-5\mathbf{g}'_2$	0	0	−8	2	−14		$\mathbf{g}''_4$	0	0	−4	1	−7
$\mathbf{g}''_4=\mathbf{g}'_4-\mathbf{g}'_2$	0	0	−4	1	−7		$\mathbf{g}''_3-2\mathbf{g}''_4$	0	0	0	0	0

El subespacio es de tres dimensiones, y una base del mismo está formada por los vectores:

$$\mathbf{g}_1, \qquad \mathbf{g}'_2=\mathbf{g}_2-2\mathbf{g}_1, \qquad \mathbf{g}''_4=\mathbf{g}_4+\mathbf{g}_1-\mathbf{g}_2.$$

[1] Una matriz es un cuadro rectangular de elementos de K, ordenados en filas y columnas.

Existe una relación de dependencia, que se escribe

$$g_3'' = 2g_4'', \qquad \text{o} \qquad 2g_1 - 3g_2 + g_3 - 2g_4 = \mathbf{O}.$$

Tres cualesquiera de los vectores dados constituyen una base del subespacio, pero, en general, la base g_1, g_2', g_4'', que verifica el criterio de independencia, es más cómoda.

OBSERVACIONES. 1.ª Cualesquiera que sean los detalles del cálculo, la matriz obtenida después de la reducción contiene r filas si es r el rango del sistema, y sus columnas verifican el criterio de independencia. Decimos que esta es una *reducida* de la matriz dada, pero entiéndase bien que esta reducida no es única (en cálculo matricial, la reducida incluye además $m-r$ filas de ceros para conservar las dimensiones).

2.ª En el caso $p_1 = 1$, $p_2 = 2$, ..., la reducida se llama *regularmente escalonada*. Presenta la siguiente forma:

$$\begin{bmatrix} a & b & c & d & e & f \\ 0 & b_1 & c_1 & d_1 & e_1 & f_1 \\ 0 & 0 & c_2 & d_2 & e_2 & f_2 \\ 0 & 0 & 0 & d_3 & e_3 & f_3 \end{bmatrix}$$

con, eventualmente, cierto número de filas de ceros.

3.ª Las operaciones elementales conducen a un proceso regular que permite obtener sistemas equivalentes a un sistema dado. A veces, resulta indicado efectuar ciertas combinaciones lineales de tipo diferente entre varios vectores del sistema. Esto es posible con la sola condición de que el nuevo sistema sea equivalente al inicial. Así, p. ej., si g_1 y g_2 pertenecen a S, y si en el sistema S' se pone en lugar de ellos $g_1 - g_2$ y $g_2 - g_1$, S' no será, en general, equivalente a S.

4.ª Examinemos el caso particular de un sistema S de n vectores de un espacio de n dimensiones y demostremos que, si constituyen una base, se puede obtener mediante operaciones elementales una matriz reducida, regularmente escalonada. No todos los vectores g pueden tener su primera coordenada nula (si no pertenecerían a un subespacio de $n-1$ dimensiones como máximo). Elegimos como g_1 uno de estos cuya primera componente no es nula, lo que eventualmente exige un cambio de dos vectores [operación *a)*, pág. 124]; se hará entonces $p_1 = 1$. No todos los vectores g' (en número de $n-1$) pueden tener su segunda componente nula, porque siéndolo ya la primera, pertenecerían entonces a un subespacio de $n-2$ dimensiones, y S no sería una base. Después de un eventual cambio se podrá hacer $p_2 = 2$, etc. La

matriz reducida obtenida se denomina triangular superior (véase número 6, Cap. 14).

EJERCICIOS:

1. Consideremos en Q^4 los sistemas siguientes:

1.º $\mathbf{a}_1=(1, 2, 0, 1)$, $\mathbf{a}_2=(2, 1, 3, 1)$, $\mathbf{a}_3=(2, 4, 0, 2)$.

2.º $\mathbf{a}_1=(2, 1, 3, -1)$, $\mathbf{a}_2=(-1, 1, -3, 1)$,
$\mathbf{a}_3=(4, 5, 3, -1)$, $\mathbf{a}_4=(1, 5, -3, 1)$.

Obténgase en cada caso una base del subespacio engendrado; hállese la dimensión.

2. Las mismas cuestiones para el sistema de Q^5:

$\mathbf{a}_1=(2, 0, 1, 3, -1)$, $\mathbf{a}_2=(1, 1, 0, -1, 1)$,
$\mathbf{a}_3=(0, -2, 1, 5, -3)$ $\mathbf{a}_4=(1, -3, 2, 9, -5)$.

3. Determínese una base de la suma y de la intersección de los subespacios engendrados por los siguientes sistemas:

$\mathbf{a}_1=(1, 2, 1, 0)$, $\mathbf{a}_2=(-1, 1, 1, 1)$;
$\mathbf{b}_1=(2, -1, 0, 1)$, $\mathbf{b}_2=(1, -1, 3, 7)$.

4. Idénticas cuestiones para los sistemas siguientes:

$\mathbf{a}_1=(1, 2, -1, -2)$, $\mathbf{a}_2=(3, 1, 1, 1)$, $\mathbf{a}_3=(-1, 0, 1, -1)$;
$\mathbf{b}_1=(2, 5, -6, -5)$, $\mathbf{b}_2=(-1, 2, -7, -3)$.

25. Aplicación a los sistemas lineales.—La resolución de los sistemas lineales por el método de eliminación es en el fondo idéntica a la reducción precedente de una matriz. La operación que reemplaza el vector $\mathbf{g}_i$ por el $\mathbf{g}'_i=\mathbf{g}_i-\lambda\mathbf{g}_{k_1}$ (véase pág. 140) es el resultado de eliminar [1] la incógnita x_{p_1} (véase el sistema [I] que sigue) entre las ecuaciones de órdenes i y k_1.

Precisemos de manera general la resolución y discusión de un sistema de m ecuaciones con n incógnitas, que se escribe:

$$\left\{\begin{array}{l} a_1^1x_1+a_1^2x_2+\ldots+a_1^nx_n=b_1, \\ a_2^1x_1+a_2^2x_2+\ldots+a_2^nx_n=b_2, \\ \quad\vdots\qquad\quad\vdots\qquad\quad\vdots \\ a_m^1x_1+a_m^2x_2+\ldots+a_m^nx_n=b_m. \end{array}\right. \qquad [I]$$

Asociamos a este sistema un sistema de m vectores del espacio

[1] Eliminación que puede hacerse por sustitución o por combinación lineal.

K^{n+1}, siendo K el cuerpo de los coeficientes de [I]. A los coeficientes de la i-ésima ecuación corresponde el i-ésimo vector $\mathbf{g}_i$,

$$\mathbf{g}_i=(a_i^1,\ a_i^2,\ \ldots,\ a_i^n,\ b_i),$$

cuya última coordenada es el término independiente. La matriz que tiene por filas estos vectores es la *matriz completa* del sistema. Recordemos que se llaman *equivalentes* dos sistemas tales que toda solución del uno es también solución del otro. Se trata, evidentemente, de una relación de equivalencia entre sistemas, de la que podemos precisar su ligazón con la teoría de los espacios vectoriales.

TEOREMA.—*Sea* $S=\{\mathbf{g}_1,\ \mathbf{g}_2,\ \ldots,\ \mathbf{g}_m\}$ *el sistema de vectores asociados al sistema de ecuaciones* [I]. *A todo sistema* S′ *equivalente a* S *está asociado un sistema de ecuaciones* [I]′ *equivalente a* [I].

Sea $\mathbf{g}'_i=\lambda_i^1\mathbf{g}_1+\lambda_i^2\mathbf{g}_2+\ldots+\lambda_i^m\mathbf{g}_m$ la expresión del i-ésimo vector de S′. Supongamos que el sistema [I] admite una solución $(x_1,\ x_2,\ \ldots,\ x_n)$ y multipliquemos la primera ecuación por λ_i^1, la segunda por λ_i^2, ..., etc. Se obtiene así la ecuación a la cual corresponde el vector $\mathbf{g}'_i$ de S′, y la solución considerada del sistema [I] satisface a dicha ecuación. La reciprocidad entre S y S′ implica que toda solución de [I]′ sea también solución de [I]. Finalmente, si [I] no tiene solución es imposible que [I]′ la tenga, en virtud de la reciprocidad anterior. Queda, pues, demostrada la equivalencia entre los sistemas [I] y [I]′.

Resolución.——Las operaciones elementales efectuadas sobre las líneas de la matriz del sistema conducen a sistemas equivalentes. Apliquemos la reducción estudiada anteriormente, pero considerando solo las n primeras columnas (coeficientes de las incógnitas) de la matriz [1]. Se puede siempre suponer (en caso necesario, basta cambiar la ordenación de las incógnitas) que la matriz reducida que así resulta es regularmente escalonada (para las r primeras columnas). Llegamos así a un sistema [I]′ equivalente a [I], y cuya matriz se escribe

$$\begin{bmatrix} c_1^1 & c_1^2 & \ldots & c_1^r & \ldots & c_1^n & d_1 \\ 0 & c_2^2 & \ldots & c_2^r & \ldots & c_2^n & d_2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & c_r^r & \ldots & c_r^n & d_r \\ 0 & 0 & \ldots & 0 & \ldots & 0 & d_{r+1} \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 0 & d_{m_r} \end{bmatrix} \qquad (\text{con } c_1^1,\ c_2^2,\ \ldots,\ c_r^r \neq 0)$$

[1] Es decir, que los números $p_1,\ p_2,\ \ldots$ de la teoría se eligen entre los n primeros enteros. Aunque, naturalmente, las operaciones se aplican a los vec-

La resolución y la discusión del sistema [I]' resultan obvios. Las $m-r$ últimas ecuaciones no contienen las incógnitas, y son necesarias las condiciones

$$d_{r+1}=d_{r+2}=\ldots=d_m=0$$

para la compatibilidad. Son además suficientes, y el sistema [I]' es indeterminado si $r<n$. Se puede, en efecto, dar valores arbitrarios a las incógnitas x_{r+1}, x_{r+2}, ..., x_n, llamadas incógnitas no principales, y resolverlo con relación a las otras.

La r-ésima ecuación determina x_r; llevando su valor a la $(r-1)$-ésima hallamos x_{r-1}, después x_{r-2}, ..., y, finalmente, x_1. Se trata de un sistema *en cascada*, cuya posibilidad queda asegurada por las condiciones c_1^1, c_2^2, ..., $c_r^r \neq 0$. Enunciemos el resultado:

TEOREMA.—*Para que un sistema de m ecuaciones con n incógnitas sea compatible, es necesario y suficiente que el rango de los vectores de las filas de la matriz completa sea igual al rango de los vectores de las filas de la matriz de los coeficientes. Si es r este rango común, pueden elegirse $n-r$ incógnitas arbitrariamente. Para $n=r$ la solución es única.*

Caso particular.—El caso $m=n$ es importante. La solución es única si $m=n=r$; es decir, si los n vectores que forman las filas de la matriz de los coeficientes son independientes en K^n.

EJEMPLO.—Con frecuencia, el método expuesto es el más adecuado para resolver un sistema numérico. Indiquemos un ejemplo, en el que el lector se encontrará con el procedimiento clásico.

A todo sistema se asocia su matriz:

$$\begin{cases} x-y+z=7 \\ 4x-2y+z=3 \\ 2x-3y+5z=2 \end{cases} \longrightarrow \begin{bmatrix} 1 & -1 & 1 & 7 \\ 4 & -2 & 1 & 3 \\ 2 & -3 & 5 & 2 \end{bmatrix}$$

Esta se reduce y se llega al sistema equivalente:

$$\begin{bmatrix} 1 & -1 & 1 & 7 \\ 0 & -1 & 3 & -12 \\ 0 & 0 & 3 & -49 \end{bmatrix} \longrightarrow \begin{cases} x-y+z=7 \\ -y+3z=-12 \\ 3z=-49. \end{cases}$$

La resolución resulta ahora inmediata.

---

tores completos. La matriz obtenida suprimiendo la última columna de la matriz completa del sistema recibe el nombre de *matriz de los coeficientes*.

Digamos para terminar que la práctica ha conducido a descubrir ciertas reglas que aumentan la rapidez del cálculo, y que se adaptan a las posibilidades de las calculadoras automáticas.

EJERCICIOS:

Resuélvanse y discútanse los siguientes sistemas:

$$\begin{cases} x_1-2x_2+3x_3-4x_4=4 \\ \quad x_2-\ x_3+\ x_4=-3 \\ x_1+3x_2\quad\ -3x_4=1 \\ \ -7x_2+3x_3+\ x_4=-3, \end{cases} \qquad \begin{cases} 2x_1+\ x_2-\ x_3-\ x_4+\ x_5=1 \\ \ x_1-\ x_2+\ x_3+\ x_4-2x_5=0 \\ 3x_1+3x_2-3x_3-3x_4+4x_5=2 \\ 4x_1+5x_2-5x_3-5x_4+7x_5=3, \end{cases}$$

$$\begin{cases} \ x_1+2x_2\quad\ -\ 3x_4+2x_5=1 \\ \ x_1-\ x_2-3x_3+\ \ x_4-3x_5=2 \\ 2x_1-3x_2+4x_3-\ 5x_4+2x_5=7 \\ 9x_1-9x_2+6x_3-16x_4+2x_5=25, \end{cases} \qquad \begin{cases} \ x_1+2x_2+3x_3-x_4=1 \\ 3x_1+2x_2+\ x_3-x_4=1 \\ 2x_1+3x_2+\ x_3+x_4=1 \\ 2x_1+2x_2+2x_3-x_4=1 \\ 5x_1+5x_2+2x_3\quad\ =2. \end{cases}$$

EJERCICIOS SOBRE EL CAPITULO VII

7-1. La dimensión de un subespacio suplementario de E_1 se llama *codimensión* de E_1. Demuéstrese la fórmula

$$\text{codim. } (E_1+E_2)+\text{codim. } (E_1\cap E_2)=\text{codim. } E_1+\text{codim. } E_2.$$

7-2. Los subespacios de dimensión 1 se suelen llamar *rectas;* los de dimensión 2, *planos* (como en geometría analítica clásica), y los de dimensión $n-1$, o de codimensión 1, hiperplanos. Demuéstrese que todo subespacio F de E es la intersección de los hiperplanos que le contienen.

7-3. Pruébese que en el espacio de los vectores con cuatro componentes reales (x_1, x_2, x_3, x_4), los vectores que verifican $x_1+x_2+x_3+x_4=0$ constituyen un subespacio de tres dimensiones. Hállese una base del mismo.

7-4. La misma cuestión para los vectores que verifican las relaciones

$$x_1+x_3=x_2-x_4=0.$$

¿Cuál es la dimensión del subespacio?

7-5. Demuéstrese que los tres vectores (2, 1, 1), (1, 3, 1), (−2, 1, 3) constituyen una base del espacio de tres dimensiones. Determínense las componentes del vector (1, 1, 1).

7-6. ¿Qué podemos decir de un sistema ligado que comprende n vectores si todo subconjunto propio, extraído del mismo, es libre?

7-7. En el espacio de los vectores de tres componentes sobre $Q(i)$, pruébese que $(1, 2i, -i)$, $(2, 1+i, 1)$, $(-1, 1, -i)$ constituyen una base. Hállense las componentes del vector (1, 2, 0).

7-8. En el espacio de cuatro dimensiones sobre Q considérense los subespacios E′ y E″ engendrados por:

E′: $(1, -2, -1, 3)$, $(0, 2, 1, -1)$, $(-2, 6, 3, -7)$;
E″: $(1, 2, 1, 1)$, $(2, 0, -1, 1)$.

Hállense las dimensiones de E′, E″, E′+E″ y E′ ∩ E″.

7-9. Si dos subespacios E′ y E″ tienen el mismo número de dimensiones y si E′ ⊂ E″, se tiene que E′=E″.

7-10. En un espacio de dimensión n se consideran dos subespacios E′ y E″ de dimensiones respectivas n' y n''. Indíquense, en cada caso, los valores máximo y mínimo de las dimensiones de E′+E″ y de E′ ∩ E″.

7-11. En el espacio E_4 (sobre Q) determínense α y β para que los tres vectores $\mathbf{V}_1=(1, 2, \alpha, 1)$, $\mathbf{V}_2=(\alpha, 1, 2, 3)$, $\mathbf{V}_3=(0, 1, \beta, 0)$ sean dependientes. Hállese la relación que existe entre ellos. Dedúzcanse las relaciones entre x_1, x_2, x_3, x_4 para que el vector correspondiente, V, pertenezca al subespacio engendrado por $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$.

7-12. Considérese la ley de recurrencia

$$f(n)=af(n-1)+bf(n-2), \qquad [1]$$

en la que los números reales a y b son dados y f designa una función desconocida que depende del entero n.

1.º Demuéstrese que el conjunto de las soluciones de [1] constituye un espacio vectorial sobre el cuerpo de los números reales.

2.º Demuéstrese que queda determinada $f(n)$ si se dan $f(0)$ y $f(1)$. Dedúzcase de ello que la solución general es

$$f(n)=\lambda f_1(n)+\mu f_2(n),$$

donde f_1 y f_2 designan dos soluciones particulares linealmente independientes, y λ y μ, dos constantes arbitrarias.

3.º Estúdiense las soluciones de [1] de la forma $f(n)=r^n$, donde r es un número real. ¿En qué caso se obtiene así la solución general de [1]?

Resuélvase $f(n)=f(n-1)+f(n-2)$ con $f(0)=f(1)=1$.

4.º Cuando existe solo una solución de la forma r^n, estúdiese la sucesión $\varphi(n)$ definida por $f(n)=r^n\varphi(n)$. Dedúzcase de ello la solución general de [1] en este caso.

7-13. Si C es una constante real, hállense las funciones derivables f, definidas sobre R y tales que

$$f'(x)=f(x+C).$$

Demuéstrese que el conjunto de estas funciones constituye un espacio vectorial. ¿Cuál es la dimensión de este espacio? Determínese una base.

PARTE II

DOS CUERPOS FUNDAMENTALES

CAPITULO VIII

EL CUERPO DE LOS NUMEROS REALES

CONSTRUCCION DE LOS NUMEROS REALES

1. Consideraciones de orden intuitivo.—El lector ha razonado y operado con «números» de los que no conocía el valor exacto: $\sqrt{2}$, π, sen 39°, log 7, etc. La validez de estos cálculos la ha considerado garantizada por el hecho de que se aproximan de manera razonable a operaciones de un tipo ideal, concebibles pero no realizables en la práctica, porque necesitan un número infinito de decimales.

Se trata, p. ej., de calcular un «número irracional» x por medio de un algoritmo. ($\sqrt{2}$ mediante el algoritmo de la raíz cuadrada, π por el procedimiento de los polígonos regulares de 2^n lados, sen 39° y log 7 por métodos desconocidos todavía del lector.) Un primer cálculo da el valor racional x_1, un segundo proporciona x_2, ..., en la n-ésima etapa obtenemos x_n, etc. Se comprende que si los x_n son aproximaciones cada vez mejores del «número» x, diferirán cada vez menos entre sí. De manera más precisa, si ϵ es un racional positivo arbitrariamente pequeño, pero fijo (p. ej., 10^{-100}), se podrá encontrar un entero fijo, N, tal que para n y p superiores a N, la diferencia $x_n - x_p$ permanezca siempre inferior a ϵ en valor absoluto. Se dice en este caso que la sucesión x_n verifica el *criterio de Cauchy*[1].

Por otra parte, el mismo «número» x puede definirse a partir de diferentes algoritmos (p. ej., $\sqrt{2}$ calculado en el sistema duodecimal; π a partir del hexágono y de polígonos de $3 \cdot 2^n$ lados), lo que conduciría a otras sucesiones $y_1, y_2, \ldots, y_n, \ldots$; $z_1, z_2, \ldots, z_n \ldots$ Se ve, pues, la necesidad de no *ligar* un «número» a *una* sucesión, sino de definir una relación de equivalencia entre las sucesiones.

Convenientemente precisadas estas nociones conducen a la construcción de los *números reales* por el método de Cantor[2].

---

[1] Augustin Louis CAUCHY (1789-1857). Ilustre matemático francés. Introdujo en análisis el rigor moderno y creó, entre otras, la teoría de las funciones de variable compleja.

[2] Georg CANTOR (1845-1918). Matemático de primer orden, creador, a pesar del sarcasmo y hostilidad de muchos de sus contemporáneos, de la teoría de conjuntos, que ha ejercido notable influencia en la matemática moderna.

En todo este capítulo el lector podrá representar los números por medio de puntos sobre una recta orientada. Se trata solo de una imagen cómoda para ayudar al razonamiento. La teoría rigurosa de la recta, en relación con los números reales, se desarrollará más adelante.

2. Sucesiones regulares. Sucesiones equivalentes.—Se llama sucesión regular y se designa por $\{x_n\}$ toda sucesión indefinida $x_1, x_2, \ldots, x_n, \ldots$ de números racionales tales que, dado un número racional ϵ arbitrariamente pequeño, pero fijo, se cumpla la desigualdad

$$|x_n - x_p| < \epsilon$$

con la condición suficiente de que n y p sean superiores a cierto número fijo N (función de ϵ); dicho de otra manera, una sucesión que verifica la condición de Cauchy.

La imagen de una sucesión regular sobre un eje Ox es una sucesión de puntos numerados, de abscisas racionales, tal que la distancia entre dos de ellos llega a hacerse tan pequeña como se quiera con la condición de tomar los índices suficientemente grandes; intuitivamente es fácil imaginarse cómo van acumulándose los puntos en torno a un punto fijo, no necesariamente racional.

El estudio que sigue se desarrollará sobre el conjunto de las sucesiones regulares. Dos sucesiones son iguales cuando lo son los elementos del mismo orden. Este criterio de igualdad se sustituirá por una equivalencia, que vamos a definir.

Sucesión equivalente a cero.—Una sucesión $\{x_n\}$ es equivalente a cero, si, dado ϵ racional arbitrariamente pequeño, la desigualdad $|x_n| < \epsilon$ se cumple mediante una condición suficiente $n > N(\epsilon)$. Con el lenguaje de la teoría de límites ello equivale a decir que x_n tiende a cero cuando n tiende a infinito. Se observa inmediatamente que una sucesión equivalente a cero es regular, porque

$$|x_n - x_p| \leqslant |x_n| + |x_p|$$

y es suficiente tomar p y n mayores que $N(\epsilon/2)$ para satisfacer la condición de Cauchy. La sucesión nula $\{0\} = 0, 0, \ldots, 0, \ldots$ es un caso particular de sucesión equivalente a cero. Indicaremos la equivalencia con la notación $\{x_n\} \sim \{0\}$.

Sucesiones equivalentes.—En general, dos sucesiones regulares, $\{x_n\}$ e $\{y_n\}$, son equivalentes [$\{x_n\} \sim \{y_n\}$] si, y solo si, $\{x_n - y_n\}$ es una

sucesión equivalente a cero. Se trata, en efecto, de una relación de equivalencia, porque:

1.° $\{x_n\} \sim \{x_n\}$, ya que $\{x_n - x_n\} = \{0\}$ es la sucesión nula;
2.° $\{x_n\} \sim \{y_n\}$ implica $\{y_n\} \sim \{x_n\}$;
3.° $\{x_n\} \sim \{y_n\}$ e $\{y_n\} \sim \{z_n\}$ implican que $\{x_n\} \sim \{z_n\}$.

En efecto, $|x_n - z_n| \leqslant |x_n - y_n| + |y_n - z_n|$. Sea ϵ arbitrario:

$$|x_n - y_n| < \frac{\epsilon}{2} \quad \text{para} \quad n > N_1, \quad \text{ya que} \quad \{x_n\} \sim \{y_n\},$$

$$|y_n - z_n| < \frac{\epsilon}{2} \quad \text{para} \quad n > N_2, \quad \text{ya que} \quad \{y_n\} \sim \{z_n\},$$

de donde $|x_n - z_n| \leqslant \epsilon$ para $n > \text{máx}\,(N_1, N_2)$.

Hemos dividido, por tanto, el conjunto de las sucesiones regulares en clases de equivalencia disjuntas (véase pág. 16) que constituyen el conjunto cociente, el cual, por definición, se llamará *conjunto de los números reales*, simbolizándolo con la letra R. El objeto de este capítulo es demostrar que las operaciones sobre sucesiones regulares inducen operaciones sobre el conjunto R, al que confieren estructura de cuerpo.

EJERCICIOS:

1. Demuéstrese que toda sucesión equivalente a una sucesión regular es regular.

2. Pruébese que si una sucesión de números racionales tiene límite sobre el cuerpo de los racionales (véase pág. 110), toda sucesión equivalente tiene el mismo límite.

3. Propiedades de las sucesiones regulares.—Citemos algunas sencillas propiedades, que serán empleadas más adelante.

Una sucesión se dice acotada superiormente en valor absoluto cuando existe un número racional, A, tal que, para todo n, $|x_n| < A$. Se puede considerar también el caso en el que la desigualdad anterior solo es válida a partir de cierto lugar, para $n > N_1$.

De manera análoga se define una sucesión acotada inferiormente.

LEMA 1.—*Toda sucesión regular está acotada superiormente en valor absoluto.*

En efecto, $|x_n - x_p| < 1$ para n y $p > N(1)$. Sea $K = N + 1$; cualquiera que sea $n > K$, tenemos:

$$|x_n| \leqslant |x_n - x_K| + |x_K| < 1 + |x_K|.$$

El mayor de los números $|x_1|$, $|x_2|$, ..., $|x_N|$, $1+|x_K|$ es una cota superior de los valores absolutos de las x_n.

LEMA 2.—*Toda sucesión regular no equivalente a cero está acotada inferiormente en valor absoluto, a partir de cierto término.*

Esto equivale a demostrar que existe un número $a>0$ y un entero N_1 tales que, para $n>N_1$, se verifica que $|x_n|>a$. Razonemos por reducción al absurdo: sea $\{x_n\}$ una sucesión para la cual la conclusión es falsa; cualesquiera que sean $a>0$ y N_1 existe al menos un $p>N_1$ tal que $|x_p|<a$. Pero existe $N(a)$ tal que n y $q>N(a)$ implican $|x_n-x_q|<a$. Elijamos $N_1=N(a)$ y $q=p$; se tiene entonces para todo $n>N(a)$,

$$\begin{aligned}|x_n| &\leqslant |x_p|+|x_n-x_p|,\\ &< a+a=2a.\end{aligned}$$

Como el número racional a es arbitrario, lo podemos elegir tan pequeño como queramos y la sucesión x_n es equivalente a cero, lo que contradice la hipótesis.

LEMA 3.—*Se obtiene una sucesión regular* $\{\xi_n\}$ *equivalente a una sucesión dada* $\{x_n\}$ *suprimiendo en ella un número* K *arbitrario, pero fijo y finito, de términos iniciales.*

En efecto, $\{\xi_n\}=\{x_{K+n}\}$ es regular y, por otra parte,

$$\{\xi_n-x_n\}=\{x_{K+n}-x_n\}$$

es equivalente a cero por la regularidad de $\{x_n\}$.

4. Grupo aditivo de R. SUMA.—Por definición, la suma de dos sucesiones regulares $\{x_n\}$ e $\{y_n\}$ es la sucesión $\{s_n\}=\{x_n+y_n\}$.

1.° La sucesión suma es regular, porque

$$|s_p-s_n|=|x_p-x_n+y_p-y_n|\leqslant |x_p-x_n|+|y_p-y_n|.$$

Para que se verifique $|s_p-s_n|<\epsilon$ es suficiente que n y p sean mayores que $N_1(\epsilon/2)$ y $N_2(\epsilon/2)$ y, por consiguiente, superiores al máximo N de N_1 y N_2.

2.° Se obtiene una sucesión $\{\sigma_n\}$ equivalente a $\{s_n\}$ si $\{x_n\}$ e $\{y_n\}$ se reemplazan por sucesiones equivalentes $\{\xi_n\}$ y $\{\eta_n\}$.

Sea $|\delta_n|=|s_n-\sigma_n|\leqslant |x_n-\xi_n|+|y_n-\eta_n|$; para que sea $|\delta_n|<\epsilon$, basta que $n>N_1(\epsilon/2)$ y $n>N_2(\epsilon/2)$; por consiguiente, $n>N(\epsilon)=$ $=$ máx (N_1, N_2).

Sucesión opuesta. Diferencia de dos sucesiones.—Por definición, la opuesta de $\{x_n\}$ es $\{-x_n\}$; es regular y queda sustituida por una equivalente al mismo tiempo que $\{x_n\}$. La diferencia se define por la adición de la opuesta; la operación es estable para la equivalencia.

CONCLUSIÓN.—El conjunto de las sucesiones regulares constituye un grupo abeliano con relación a la adición, ya que la asociatividad y la conmutatividad son evidentes; el elemento neutro $\{0\}$ es regular, así como el opuesto de todo elemento. Si de las sucesiones regulares se pasa a las clases de equivalencia, la estabilidad de la ley estudiada induce en estas una adición que posee las mismas propiedades. La estabilidad del elemento opuesto completa la existencia de un grupo aditivo sobre el conjunto R.

5. Grupo multiplicativo de R. Producto de dos sucesiones regulares. Por definición, las sucesiones regulares $\{x_n\}$ e $\{y_n\}$ tienen por producto la sucesión $\{p_n\} = \{x_n y_n\}$.

1.° $\{p_n\}$ es regular. En efecto,

$$|x_q y_q - x_n y_n| = |x_q(y_q - y_n) + y_n(x_q - x_n)|$$

y, si A y B son cotas superiores (lema 1) de x_n e y_n, respectivamente, tendremos:

$$|p_q - p_n| < A|y_q - y_n| + B|x_q - x_n|.$$

Para lograr que esta diferencia sea menor que ϵ, haremos $|y_q - y_n| < \epsilon/2A$ con q y n superiores a $N_1(\epsilon/2A)$ y $|x_q - x_n| < \epsilon/2B$ con q y n superiores a $N_2(\epsilon/2B)$; condición suficiente será que q y n sean superiores a $N = \text{máx}(N_1, N_2)$.

2.° Se obtiene una sucesión equivalente $\{\pi_n\}$ reemplazando $\{x_n\}$ e $\{y_n\}$ por otras equivalentes $\{\xi_n\}$ y $\{\eta_n\}$.

En efecto,

$$|p_n - \pi_n| = |x_n(y_n - \eta_n) + \eta_n(x_n - \xi_n)|.$$

La demostración se termina como anteriormente, introduciendo las cotas superiores A y B de $|x_n|$ e $|y_n|$.

3.° El producto de dos sucesiones no equivalentes a cero no es equivalente a cero. En efecto, según el lema 2, existe un número $a \neq 0$ tal que $|x_n| > a$ e $|y_n| > a$ para $n > N_1$. Se deduce de ello que $|x_n y_n| > a^2$, de donde $\{p_n\}$ no es equivalente a cero.

De lo anterior resulta que, si el producto de dos sucesiones es equivalente a cero, también lo será una de ellas al menos.

Inversa de una sucesión regular.—Sea una sucesión $\{x_n\}$ regular y no equivalente a cero: $|x_n| > a$ para $n > N_1$. Definimos una nueva sucesión y_n tal que

$$y_n = 0 \qquad \text{para} \qquad n \leqslant N_1,$$
$$y_n = 1/x_n \qquad \text{para} \qquad n > N_1.$$

La sucesión producto $\{x_n y_n\}$ tiene sus N_1 primeros términos nulos, siendo los otros iguales a 1; según el lema 3, $\{x_n y_n\} \sim \{1\}$.

Por definición, la sucesión $\{y_n\}$ se llama inversa de la $\{x_n\}$.

1.° La inversa de una sucesión regular y no equivalente a cero posee estos mismos caracteres.

Se tendrá $|x_n - x_p| < \epsilon a^2$ para n y $p > N$. Si se eligen también n y p superiores a N_1, tendremos $|x_n x_p| > a^2$. Por consiguiente,

$$\frac{|x_n - x_p|}{|x_n x_p|} = \left| \frac{1}{x_p} - \frac{1}{x_n} \right| < \epsilon.$$

Por tanto, la sucesión $\{y_n\}$ es regular, pero no equivalente a cero, porque en tal caso lo sería el producto $\{x_n y_n\}$ (pág. 155) y no podría ser equivalente a 1.

2.° La inversa es estable respecto de la relación de equivalencia.

Sea $\{\xi_n\} \sim \{x_n\}$, y η_n su inversa. Se tiene:

$$|y_n - \eta_n| = \frac{|\xi_n - x_n|}{|x_n \xi_n|}$$

y, para n suficientemente grande, el segundo miembro será tan pequeño como se desee, ya que $|x_n| > a$ y $|\xi_n| > a'$.

CONCLUSIÓN.—El producto de sucesiones regulares es una ley de composición estable respecto de la relación de equivalencia, lo que define un producto sobre el conjunto R: la operación es asociativa y conmutativa; el elemento neutro es la clase de las sucesiones equivalentes a $\{1\}$, y todo elemento no nulo de R (correspondiente a sucesiones no equivalentes a cero) admite un inverso.

En resumen: los elementos no nulos de R constituyen un grupo multiplicativo [1].

6. El cuerpo de los números reales.—Para demostrar que R constituye un cuerpo conmutativo es suficiente comprobar que el producto

[1] Sobre el conjunto de las sucesiones regulares los elementos no nulos no constituyen un grupo multiplicativo, porque los elementos equivalentes a cero (y distintos de $\{0\}$) carecen de inversos.

es distributivo con relación a la suma; es decir,

$$[\{x_n\} + \{y_n\}] \times \{z_n\} = \{x_n\} \times \{z_n\} + \{y_n\} \times \{z_n\}.$$

La demostración es inmediata.

Designaremos ahora por una letra cualquier elemento de R (X en lugar de $\{x_n\}$) y podremos hablar de estos elementos como de los números sin repetir necesariamente el hecho de que X representa una clase de sucesiones regulares equivalentes.

Demostremos que R contiene un subcuerpo isomorfo al cuerpo Q de los números racionales. Para ello es suficiente considerar las sucesiones $\{x_n\} = \{p/q\}$; es decir, $x_n = p/q$ cualquiera que sea n. Si a toda sucesión asociamos el racional correspondiente, existirá una correspondencia biunívoca, que se conserva en las dos leyes de composición, adición y multiplicación. Las sucesiones consideradas forman, pues, un subcuerpo isomorfo a Q, que, siguiendo un método repetidamente utilizado, identificaremos con el cuerpo de los números racionales.

TEOREMA.—*El cuerpo de los números reales tiene característica infinita.*

Ya que para un elemento cualquiera X, diferente de cero, la suma

$$X + X + \ldots + X = p \times X$$

jamás es nula.

PROPIEDADES FUNDAMENTALES DEL CUERPO DE LOS NUMEROS REALES

7. R es un cuerpo ordenado. LEMA.—*En una sucesión regular no equivalente a cero, todos los términos, a partir de uno de ellos, tienen el mismo signo.*

Es suficiente para ello que los términos positivos o los negativos estén en número limitado. Supongamos que hay un número infinito de términos positivos,

$$x_{p_1}, \ldots, x_{p_2}, \ldots, x_{p_3}, \ldots, x_{p_i}, \ldots$$

y un número infinito de términos negativos,

$$x_{n_1}, \ldots, x_{n_2}, \ldots, x_{n_3}, \ldots, x_{n_j}, \ldots;$$

en tal hipótesis, sería posible hacer $|x_{p_i} - x_{n_j}| < \epsilon$ para p_i y $n_j > N(\epsilon)$; de donde, x_{p_i} y $|x_{n_j}| < \epsilon$ y la sucesión sería equivalente a cero.

En estas condiciones, se demuestra que se cumplen los axiomas de orden sobre un cuerpo. Hemos definido ya las sucesiones nulas, y el lema anterior permite definir las positivas y negativas. Es obvio que toda sucesión equivalente a una sucesión positiva (o negativa) es ella misma positiva (o negativa). Además, una sucesión es o positiva o negativa o nula, y las sucesiones positivas constituyen un conjunto cerrado con relación a la suma y al producto: R es un cuerpo ordenado. Señalemos que el orden, de esta manera introducido, es compatible con el orden definido anteriormente sobre el cuerpo de los números racionales. Un número real $\{x_n\}$ es mayor que $\{y_n\}$ cuando $\{x_n - y_n\} > 0$. Las reglas de cálculo con desigualdades demostradas (pág. 79) para los anillos ordenados y completadas en la página 108 son también válidas aquí.

Observemos finalmente que, en todo cuerpo ordenado, entre dos elementos $X < Y$ existe una infinidad de otros: p. ej.,

$$X < \frac{X+Y}{2} < Y.$$

Valor absoluto.—La noción de valor absoluto se extiende a R: $|X|$ designa el elemento positivo del par X, $-X$, cuando $X \neq 0$; las propiedades son las mismas que en Z o en Q:

$$|XY| = |X| \times |Y| \qquad \text{y} \qquad |X+Y| \leqslant |X| + |Y|.$$

Límite.—El concepto de límite coincide con el ya definido sobre Q (pág. 110). Si es X_n una sucesión de números reales[1] (uno para cada valor del entero n), se dice que X_n tiende hacia A para n infinito, cuando a todo número real ϵ, por pequeño que sea, se puede hacer corresponder un entero N tal que $n > N$ implica $|X_n - A| < \epsilon$.

Prácticamente, si X_n está definida por una sucesión regular de números racionales $\{x_{n,m}\}$ (donde m indica el orden en una sucesión dada), A lo está por la sucesión $\{a_n\}$ y si se da ϵ racional, debe ser siempre posible determinar $N(\epsilon)$ tal que eligiendo $n > N(\epsilon)$ podamos hallar en la sucesión $\{x_{n,m}\}$ correspondiente un término $x_{n,M}$ (cuyo lugar $M(\epsilon, n)$ se deduce a partir de ϵ y n) tal que todos los términos siguientes ($m > M$) cumplan la condición

$$|x_{n,m} - a_m| < \epsilon.$$

[1] Es decir, una sucesión en la que cada elemento es a su vez una sucesión regular de números racionales.

Recíprocamente, si esta condición queda satisfecha, de ella se deduce que $|X_n - A| \leqslant \epsilon$; el signo de igualdad carece de importancia ya que ϵ es arbitrario.

Se obtiene una aplicación inmediata tomando una sucesión X_n de números racionales formada a partir de una sucesión de Cauchy $X = \{x_n\}$; sobre R, X_n será, por consiguiente, la sucesión $\{x_n, x_n, \ldots, x_n, \ldots\}$. Cuando n tiende a infinito, X_n tiende hacia X. En efecto, $X - X_n$ está definida por la sucesión

$$x_1 - x_n, \quad x_2 - x_n \quad \ldots, \quad 0, \quad x_{n+1} - x_n, \quad \ldots \quad x_p - x_n, \quad \ldots$$

cuyo término general satisface a $|x_p - x_n| < \epsilon$, para n y p superiores a $N(\epsilon)$, ya que la sucesión inicial es regular. Se elegirá, pues, un n particular y, cualquiera que sea $p > N(\epsilon)$, quedará satisfecha la condición, de donde $|X_n - X| \leqslant \epsilon$.

Este resultado permite comprender mejor el objeto perseguido con la construcción de los números reales: el cuerpo R es tal que una sucesión regular de números racionales, que en general carece de límite en Q, lo admite siempre sobre R, y dicho límite es el mismo elemento definido por la sucesión [1].

EJERCICIOS:

1. Demuéstrese que toda sucesión equivalente a una sucesión positiva es positiva.

2. Pruébese que el orden definido sobre el cuerpo R es compatible con el ya establecido sobre el cuerpo de los números racionales.

3. Sobre un cuerpo ordenado (Q o R) si la sucesión X_n tiende hacia A y la Y_n hacia B, la sucesión X_nY_n tiende hacia AB. Si $B \neq 0$, la sucesión X_n/Y_n tiende hacia A/B.

8. R es un cuerpo arquimediano.—Esto significa que si a y b son dos números reales que verifican la desigualdad $0 < a < b$, existen siempre números enteros p tales que $b < pa$.

Supongamos a y b definidos por las sucesiones $\{a_n\}$ y $\{b_n\}$, y recordemos que una sucesión regular $\{b_n\}$ está acotada superiormente por un número racional β y que una sucesión regular positiva $\{a_n\}$ lo está inferiormente; la relación $|a_n| > \alpha$ se convierte en este caso en $a_n > \alpha$, con α racional. Por ser el cuerpo Q arquimediano será posible hallar un número p entero tal que $p\alpha > \beta$. Por otra parte, la

[1] De igual forma que la construcción de Q ha permitido definir un cociente para todo par a, $b (b \neq 0)$ de elementos de Z, siendo este cociente el elemento de Q asociado al par (a, b) considerado.

relación $a_n > \alpha$ implica $a \geqslant \alpha$, de donde $pa \geqslant p\alpha$. De igual forma, se tiene que $b \leqslant \beta$ y, en consecuencia, $pa > b$.

Aplicación. *Valor aproximado de un número real en menos de* ϵ. Si consideramos la progresión

$$\ldots -2\epsilon, \ -\epsilon, \ 0, \ \epsilon, \ 2\epsilon, \ 3\epsilon, \ \ldots, \ n\epsilon, \ \ldots,$$

en virtud del teorema anterior, cualquier número x estará comprendido entre dos términos de la sucesión, o bien confundido con uno de ellos. Tenemos, pues,

$$n\epsilon \leqslant x < (n+1)\,\epsilon,$$

donde n puede ser positivo o negativo.

El número $n\epsilon$ se llama valor aproximado por defecto de x en menos de ϵ, y el número $(n+1)\epsilon$ es su valor aproximado por exceso. Prácticamente, se toma para ϵ un número racional de la forma 10^{-p}.

9. R es un cuerpo completo.—El cuerpo R se ha obtenido a partir de Q estudiando las sucesiones regulares de elementos de Q. Cabe pensar inmediatamente en reiterar el procedimiento y definir sucesiones regulares de números reales; es decir, sucesiones $\{X_n\}$ tales que

$$|X_n - X_p| < \epsilon, \qquad \text{para} \qquad n > N(\epsilon) \ \text{y} \ p > N(\epsilon),$$

siendo ϵ un número real arbitrariamente pequeño. En particular, dos sucesiones $\{X_n\}$ e $\{Y_n\}$ serán equivalentes cuando la sucesión $\{X_n - Y_n\}$ sea equivalente a cero; esto es, cuando $|X_n - Y_n| < \epsilon_1$ para $n > N_1(\epsilon_1)$, donde ϵ_1 es arbitrariamente pequeño. ¿Obtendremos de esta manera un supercuerpo de R? El siguiente teorema demuestra lo contrario.

TEOREMA.—*Sobre el cuerpo* R, *toda sucesión regular de elementos reales es equivalente a una sucesión regular de elementos racionales, y admite como límite un elemento de* R.

Sea X_n un elemento cualquiera de tal sucesión; el número real X_n está definido por una sucesión regular $\{x_{n,\lambda}\}$ de números racionales que tienden hacia X_n cuando el índice λ tiende a infinito. Se puede, pues, elegir el lugar λ de manera que

$$|X_n - x_{n,\lambda}| < 10^{-n}.$$

Elegido λ de esta forma para cada valor de n, hagamos $x_{n,\lambda} = y_n$. Sobre el cuerpo de los números reales la sucesión $\{X_n\}$ es equivalente a

la sucesión $\{y_n\}$, porque $\{X_n - y_n\}$ es una sucesión equivalente a cero. Por ser la sucesión $\{X_n\}$ regular, también lo es la $\{y_n\}$ y define un elemento X de R; $\{X_n\}$, que es equivalente a ella, admite el mismo límite [1].

Este resultado se expresa diciendo que R es completo. Cabe también enunciar la propiedad de la forma siguiente:

TEOREMA.—*Una condición necesaria y suficiente para que una sucesión de números reales tienda hacia un límite es que sea regular, es decir, que verifique el criterio de Cauchy.*

10. Principio de los intervalos.—Se llama intervalo cerrado (a, b) al conjunto de números reales x tales que $a \leqslant x \leqslant b$; los números a y b son sus extremos, y $b - a$, su longitud [2]. Un segundo intervalo cerrado (a', b'), tal que $a \leqslant a' < b' \leqslant b$, se dice que está contenido en (a, b); también se emplea la expresión *intervalos encajados.*

TEOREMA.—*Si tenemos una sucesión indefinida de intervalos cerrados* I_n *tales que* I_{n+1} *esté contenido en* I_n *y que la longitud de* I_n *tienda a cero cuando n tiende a infinito, existe un número y uno solo común a todos los intervalos.*

En efecto, sean a_1 y b_1, a_2 y b_2, ..., a_n y b_n, ... los extremos de los intervalos de la sucesión; puesto que son intervalos encajados, se tendrá siempre, para $p > n$,

$$a_n \leqslant a_p < b_p \leqslant b_n,$$

es decir, que todo número de I_p pertenece a I_n. De ello resulta que la sucesión de números reales

$$a_1, b_1, a_2, b_2, \ldots, a_n, b_n, \ldots$$

es regular: a todo ϵ se puede asociar un N tal que $b_n - a_n < \epsilon$ si $n \geqslant N$, de donde, para n y p superiores a N, se verifica la desigualdad $|x_n - x_p| < \epsilon$ ($x = a$ o b). De esta manera queda definido un número X común a todos los intervalos, ya que, cualquiera que sea n, $X - a_n$ es una sucesión positiva y $X - b_n$ una sucesión negativa. No puede existir otro número X' que goce de la misma propiedad, porque $b_n - a_n > |X - X'|$ y el primer miembro no podría tender hacia cero.

[1] Véanse ejercicios 1 y 2, pág. 153. La demostración hecha sobre el cuerpo de los números racionales es válida para el cuerpo de los números reales.

[2] Se observará el lenguaje geométrico, que quedará plenamente justificado después del estudio de la medida de longitudes sobre un eje.

Este teorema es importante por su correspondencia con la geometría (véase Cap. 9, axioma de Cantor-Dedekind) y también debido a su carácter a la vez intuitivo y manejable.

Aplicación.—Se dice que un conjunto E de números reales posee un número mayorante o cota superior A (respectivamente, número minorante o cota inferior) cuando $x < A$ (respectivamente, $x > A$) para todo $x \in E$.

Se llama *límite superior* de un conjunto de números reales E, un número a tal que, para todo $x \in E$, se verifique $x \leqslant a$ y que si $a' < a$, exista al menos un x superior a a'. De manera análoga se define el límite inferior.

TEOREMA.—*Todo conjunto acotado superiormente (inferiormente) de números reales posee un límite superior (inferior).*

Sea A la cota superior, y $A' < A$ un número inferior por lo menos a un x del conjunto. Si es A'' el punto medio de AA', pueden ocurrir dos casos: A'' es mayorante del conjunto, o A'' queda rebasado o igualado por al menos un elemento del mismo. En el primer caso reemplazamos el intervalo (A', A) por (A', A''), y en el segundo, por (A'', A); en ambos casos el intervalo elegido está contenido en (A', A) y puede desempeñar el mismo papel. Continuando de este modo se obtiene una sucesión indefinida de segmentos encajados, cuya longitud tiende a cero, que define un número real B, límite superior del conjunto; en efecto:

1.° no pueden existir elementos $x > B$, porque un segmento de la sucesión, de longitud inferior a $x - B$, tendría como extremo de la derecha un número de (B, x) y no podría ser mayorante;

2.° si $B' < B$, un intervalo de longitud inferior a $B - B'$ tiene el extremo de la izquierda sobre (B', B), de donde resulta la existencia al menos de un elemento del conjunto superior a B'.

COROLARIO.—*Toda sucesión creciente y acotada superiormente tiende hacia un límite.*

Existe un límite superior B y, visto que es creciente, si x_p es uno de los términos de la sucesión que está en el intervalo $(B - \epsilon, B)$, estarán también en ese intervalo todos los términos posteriores a x_p; por consiguiente, B es el límite de la sucesión.

UNA PROPIEDAD FUNDAMENTAL DE LAS FUNCIONES CONTINUAS

11. Definición.—Una función $x \longrightarrow f(x)$, definida sobre un intervalo (a, b) del cuerpo de los números reales y que toma sus valores en dicho cuerpo, se dice continua para el valor $x = x_0$ del intervalo, cuando la diferencia $|f(x) - f(x_0)|$ puede hacerse menor que todo número real dado, positivo y arbitrariamente pequeño, con la condición suficiente de que $|x - x_0| < \eta$.

De lo anterior resulta que si una sucesión $x_1, x_2, \ldots, x_n, \ldots$ tiende hacia el número x_0, la sucesión $f(x_1), f(x_2), \ldots, f(x_n), \ldots$ tiende hacia $f(x_0)$.

12. Teorema.—*Toda función continua sobre el cuerpo de los números reales, definida en el intervalo (a, b) y tal que los valores f(a) y f(b) sean de signos contrarios, se anula al menos una vez en dicho intervalo.*

Formemos $f\left(\frac{a+b}{2}\right) = f(x_1)$. Si $f(x_1) = 0$ el teorema está demostrado; si no, reemplazamos el intervalo (a, b) por (a, x_1) o (x_1, b) según que $f(x_1)$ sea del signo de $f(b)$ o del de $f(a)$. El intervalo obtenido, de longitud $\frac{b-a}{2}$, queda contenido en (a, b) y puede desempeñar el mismo papel. Así se continúa indefinidamente, salvo que obtengamos un valor x_n raíz de $f(x) = 0$. Si no, queda construida de ese modo una sucesión indefinida de intervalos encajados cuya longitud tiende a cero, y sea X el elemento común a todos esos intervalos. En virtud de la continuidad, $f(X) = 0$. En efecto, siendo X el límite común de los extremos de la izquierda y de los extremos de la derecha de los intervalos, se tiene:

a la izquierda: $f(x_i) \leqslant 0$, de donde $\lim f(x_i) = f(X) \leqslant 0$,
a la derecha: $f(x_j) \geqslant 0$, de donde $\lim f(x_j) = f(X) \geqslant 0$.

Por consiguiente, $f(X) = 0$.

13. Caso particular.—Completemos el estudio de los polinomios por el teorema siguiente:

TEOREMA.—*Todo polinomio es una función continua sobre el cuerpo de los números reales.*

Supongamos que el valor x_0 considerado pertenece a un intervalo

(a, b), y sean $|a|$ y $|b|$ inferiores a M. Si

$$P(x) = c_0 + c_1 x + \ldots + c_n x^n,$$

un cálculo formal conduce a la identidad

$$|P(x) - P(x_0)| = |x - x_0| \times |c_1 + c_2(x_0 + x) + c_3(x_0^2 + x_0 x + x^2) + \ldots + c_n(x_0^{n-1} + x x_0^{n-2} + \ldots + x^{n-1})|,$$

$$|P(x) - P(x_0)| < |x - x_0|[|c_1| + 2M|c_2| + 3M^2|c_3| + \ldots + nM^n|c_n|] = |x - x_0| \times K;$$

el número K no depende ni de x, ni de x_0 en el interior de (a, b). Es suficiente, pues, elegir $|x - x_0| < \eta = \dfrac{\epsilon}{K}$ para que se cumpla la desigualdad $|P(x) - P(x_0)| < \epsilon$.

OBSERVACIÓN.—Cuando el valor obtenido para η es independiente del número x_0 de (a, b) se dice que la continuidad es uniforme.

EJEMPLO.—Un polinomio de grado impar tiene al menos una raíz. En efecto, se demuestra sin dificultad que existen dos valores de x, x_1 y x_2, de signos contrarios y de valores absolutos suficientemente grandes para que $f(x_1) \cdot f(x_2) < 0$.

14. Raíz *n*-ésima de un número real. TEOREMA.—*Si A es positivo, la ecuación* $x^n - A = 0$ *admite una raíz positiva y una sola.*

La ecuación $x^n - A = 0$ es tal que $P(0) < 0$, y $P(A) > 0$ (suponemos $n \geqslant 2$) para $A > 1$, y $P(1) > 0$ si $A < 1$. Existe, pues, una raíz positiva al menos. No podría haber varias porque la condición $x > x'$ implica para números positivos $x^n > x'^n$.

El número real definido de este modo, raíz n-ésima de A, se escribe $x = \sqrt[n]{A} = A^{1/n}$. De manera más general hacemos $x^p = A^{p/n}$, o lo que es lo mismo, $x^p = \sqrt[n]{A^p}$. De donde se deducen las reglas de cálculo con exponentes fraccionarios:

$$A^{p/n} \times B^{p/n} = (AB)^{p/n},$$

$$A^{p/n} \times A^{p'/n'} = A^{(pn' + p'n)/nn'},$$

$$(A^{p/n})^{p'/n'} = A^{pp'/nn'}.$$

EJERCICIOS:

1. Simplifíquese la fracción

$$\frac{15a^{7/10}-10a^{1/5}b^{1/6}-9a^{1/2}b^{1/4}-6b^{5/12}}{5a^{1/5}-3b^{1/4}}.$$

Resp.: $3a^{1/2}-2b^{1/6}$.

2. Resuélvase la ecuación en x

$$a^{-1/3}(x+1)=a^{2/3}-x(a^{1/3}+1).$$

Resp.: $x=\sqrt[4]{a}-1$.

3. Hállese el límite, cuando x tiende a infinito, de

$$\frac{\sqrt[4]{x+1}-\sqrt[4]{x}}{\sqrt[3]{x+1}-\sqrt[3]{x}}\cdot\frac{\sqrt[3]{x}}{\sqrt[4]{x}}$$

4. Demuéstrese que $\operatorname{tag}\theta=\left(\frac{b}{a}\right)^{1/3}$ implica

$$\frac{a}{\cos\theta}+\frac{b}{\operatorname{sen}\theta}=(a^{2/3}+b^{2/3})^{3/2}.$$

EJERCICIOS SOBRE EL CAPITULO VIII

8-1. Designando con a un número positivo, se considera la sucesión

$$x_1=\sqrt{a},\qquad x_2=\sqrt{a+x_1},\qquad \ldots,\qquad x_n=\sqrt{a+x_{n-1}}\qquad \ldots$$

Demuéstrese que es una sucesión creciente acotada superiormente. ¿Cuál es su límite?

8-2. Designando con a un número positivo, sea la sucesión

$$x_1=\frac{a}{x_1^2},\qquad x_2=\frac{a}{x_2^2},\qquad \ldots,\qquad x_n=\frac{a}{x_{n-1}^2},\qquad \ldots$$

Exprésese x_n en función de a y de x_1 (resultan dos casos según la paridad de n). Dedúzcase de lo anterior que una de las sucesiones $x_1, x_3, \ldots, x_{2p+1}, \ldots$ y $x_2, x_4, \ldots, x_{2p}, \ldots$, tiende hacia cero y la otra hacia infinito.

8-3. Las dos relaciones de recurrencia

$$a_{n+1}=\sqrt{a_nb_n}\qquad y\qquad b_{n+1}=\frac{a_n+b_n}{2}$$

permiten, partiendo de $a_0=a$ y $b_0=b$ dados $(0<a<b)$, calcular sucesivamente a_1 y b_1, a_2 y b_2, a_3 y b_3, etc.

Pruébese que la sucesión a_n es creciente y acotada superiormente y la sucesión b_n es decreciente y acotada inferiormente, y que $a_n<b_n$. Demués-

trese que estas dos sucesiones tienen un límite común l (llamado media aritmético-geométrica de a y b).

8-4. Sea la ecuación $x^n+x^{n-1}+\ldots+x^2+x-1=0$. Pruébese que para todo valor de n admite una raíz real positiva y solo una, ξ_n.

Demuéstrese que las ξ_n forman una sucesión monótona decreciente y calcúlese su límite.

8-5. Sea la sucesión $S_n=1+1/2+\ldots+1/n$. Si a designa un número entero fijo, demuéstrese que la sucesión $S_{na}-S_n$ es monótona, creciente y acotada superiormente. Sea $\lambda(a)$ su límite.

Si a y b son dos enteros, pruébese que $S_{na}-S_{nb}$ admite un límite igual a $\lambda(a)-\lambda(b)$. Si c es otro entero demuéstrese que $\lambda(ac)-\lambda(bc)=\lambda(a)-\lambda(b)$. Dedúzcase de lo anterior que $\lambda(a)-\lambda(b)$ es de la forma $\mu\left(\frac{a}{b}\right)$. Pruébese que

$$\mu\left(\frac{a}{b}\right)+\mu\left(\frac{a'}{b'}\right)=\mu\left(\frac{aa'}{bb'}\right).$$

8-6. ¿Qué condición deben cumplir los números racionales a, b, c, d para que $\dfrac{a+bx}{c+dx}$ sea racional, cualquiera que sea el número real x?

8-7. Dado un número real $a>0$, demuéstrese que es posible hallar un entero n tal que $1/n<a$ (empléese el orden arquimediano). Dedúzcase de ello que entre dos números reales cualesquiera, a y b, se puede encontrar siempre un número racional m/n.

8-8. A una expresión decimal de infinitas cifras $a,a_1a_2\ldots a_n\ldots$, en la que las letras a_1, a_2, ..., a_n, ... designan dígitos y a es un número entero, asociamos la sucesión $x_0=a$, $x_1=a,a_1$, ..., $x_n=a,a_1a_2\ldots a_n$ (número decimal). Demuéstrese que la sucesión x_n es regular. Se dice que la expresión decimal considerada define un número real.

8-9. Si las sucesiones regulares asociadas a dos expresiones decimales de infinitas cifras son equivalentes, estas expresiones son, en general, idénticas. La única excepción se obtiene cuando una cifra de una de las expresiones es inferior en una unidad a la cifra correspondiente de la otra, y la primera está seguida únicamente por nueves y la segunda por ceros.

8-10. Demuéstrese que toda sucesión regular, x_n, es equivalente a una sucesión regular decimal correspondiente a una expresión decimal de infinitas cifras. Se asociará a x_n la sucesión u_n definida por la doble desigualdad

$$u_n=\frac{\alpha}{10^n}\leqslant x_n<\frac{\alpha+1}{10^{n+1}}\qquad(\alpha\text{ entero})$$

y se demostrará que en la sucesión de las u_n existe un lugar r_1 a partir del cual todas comienzan por la misma cifra, y un lugar r_2 a partir del cual todas empiezan por las dos mismas cifras, etc. Demuéstrese que $\{x_n\}$ es equivalente a $\{u_{r_i}\}$.

8-11. Pruébese que el método corriente de extraer la raíz cuadrada de un número racional o real (representado por una expresión decimal de infinitas cifras) define un número real cuyo cuadrado es igual al número dado.

8-12. Demuéstrese que el conjunto de los números reales, x, comprendidos entre 0 y 1, no es numerable (véase pág. 20). Hemos visto que x no admite más que una representación de la forma $x=0,a_1a_2\ldots a_n\ldots$, si excluimos las sucesiones formadas únicamente por 9 a partir de cierta cifra. Enumerar las x equivale a establecer una correspondencia biunívoca entre su conjunto y el de los enteros naturales. Se tendría,

$$\begin{array}{l} x_1=0,a_{11}a_{12}a_{13}\ldots a_{1n}\ldots \\ x_2=0,a_{21}a_{22}a_{23}\ldots a_{2n}\ldots \\ \vdots \\ x_n=0,a_{n1}a_{n2}a_{n3}\ldots a_{nn}\ldots \\ \vdots \end{array}$$

Hagamos $b_i=a_{ii}+1$ si $a_{ii}=0$ ó 1 ó 2 ... ó 7 y $b_i=7$ si $a_{ii}=8$ ó 9. Demuéstrese que el número $0,b_1b_2\ldots b_n\ldots$ no figura entre los x.

8-13. *Números p-ádicos.*—Observemos que todas las propiedades de las sucesiones regulares están basadas en las propiedades de la función valor absoluto, a saber

$$|a|=0 \quad \text{para} \quad a=0, \qquad |a+b|\leqslant|a|+|b| \quad \text{y} \quad |ab|=|a|\cdot|b|.$$

Sea p un número primo fijo; todo racional x puede escribirse en la forma $x=\dfrac{a}{b}p^{\lambda}$, donde λ designa un entero relativo. Hacemos $|[x]|=p^{-\lambda}$; demuéstrese que la función $|[x]|$ tiene las mismas propiedades que $|x|$.

1.° Precísense las nociones de límite y de sucesión regular de racionales empleando $|[x]|$ en lugar de $|x|$.

2.° Sígase lo hecho a partir de Q para obtener un cuerpo completo reemplazando $|x|$ por $|[x]|$. Se obtendrá de esta forma el cuerpo completo de los números p-ádicos.

3.° Demuéstrese que todo número p-ádico puede representarse por un desarrollo finito o infinito, $a_0+a_1p+\ldots+a_np^n+\ldots$, en el que el entero a_n satisface a $0\leqslant a_n<p$. Enúnciense las reglas de cálculo para obtener la suma y el producto de dos de las expresiones anteriores.

8-14. Sean α y β dos números reales positivos. Si m y n son dos enteros naturales, representemos por $[\alpha m]$ y $[\beta n]$ las partes enteras de los números αm y βn. Suponemos que, por describir m y n el conjunto N* de los enteros naturales, los valores de $[\alpha m]$ y $[\beta n]$ reunidos en un mismo conjunto describen igualmente N* sin laguna ni repetición. Pruébese que esto solo es posible para α y β irracionales y ligados por la relación $\dfrac{1}{\alpha}+\dfrac{1}{\beta}=1$.

CAPITULO IX

ISOMORFISMO FUNDAMENTAL EN EL CUERPO DE LOS NUMEROS REALES

COMPARACION ENTRE EL GRUPO ADITIVO DE LOS ELEMENTOS DE R Y EL GRUPO MULTIPLICATIVO DE SUS ELEMENTOS POSITIVOS

Hemos construido y estudiado el conjunto R de los números reales en cuanto cuerpo. Para profundizar este estudio vamos a desgajar de R el grupo R_A obtenido considerando la adición, y el grupo R_M^+ que resulta cuando se consideran los elementos positivos con la multiplicación como ley de composición.

Siempre que hablemos de orden en estos grupos, se tratará del orden inducido por el orden de R.

1. Propiedades del grupo aditivo.—R_A posee, entre otras, las propiedades siguientes:

A_1) *Es un grupo abeliano dotado de una relación de orden compatible con la ley de grupo.*

Dicho de otra manera, $x \leqslant y$ implica $x+z \leqslant y+z$.

A_2) *Entre dos elementos de* R_A *existen infinitos otros.*

Esto es consecuencia del orden en R.

B_1) R_A *es arquimediano.*

Dicho de otra manera, cualesquiera que sean a y b tales que $0 < a < b$, la sucesión $a_1 = a$, $a_2 = a + a$, ..., $a_n = a + a + \ldots + a = na$, ... no está acotada superiormente y acaba por sobrepasar a b.

Esto resulta también del orden arquimediano en R.

B_2) *Es aplicable el principio de los intervalos.*

Ya que dicho principio se aplica en R, que tiene los mismos elementos que R_A.

2. Propiedades del grupo multiplicativo R_M^+.—Vamos a demostrar que posee, entre otras, las propiedades A_1, A_2, B_1, B_2.

A_1 resulta como antes del hecho de estar R_M^+ sumergido en el cuerpo ordenado R.

La escritura difiere solo ligeramente de la que empleamos antes; ahora $x \leqslant y$ implica $xz \leqslant yz$ para todo $z \in R_M^+$.

A_2 es exactamente la propiedad enunciada en la página anterior.

B_1 necesita demostración. El orden de un grupo abeliano denotado multiplicativamente es arquimediano si, cualesquiera que sean a y b tales que $1 < a < b$, la sucesión $a_1 = a$, $a_2 = a \cdot a = a^2$, $a_3 = a^3$, ..., $a_n = a^n$ ... no está acotada superiormente y termina, por tanto, sobrepasando a b.

Sea, pues, $a = 1 + \alpha$, $\alpha > 0$; sabemos que $(1+\alpha)^n > 1 + n\alpha$, y bastará, por consiguiente, tomar $1 + n\alpha > b$, es decir, $n > \dfrac{b-1}{\alpha}$.

Deducimos de lo anterior que la sucesión $a^{1/n}$ tiende a 1 cuando el entero n se hace infinito. En efecto, la sucesión es decreciente y acotada inferiormente por 1. Si $a^{1/n}$ se mantuviese siempre superior a b, o sea $b < a^{1/n}$, se tendría, cualquiera que fuera n, $b^n < a$, en contra del orden arquimediano. El mismo resultado para $a < 1$, ya que la sucesión es creciente y está acotada superiormente por 1.

B_2: el principio de los intervalos se aplica en R_M^+. Debe observarse, sin embargo, que es necesario un cambio de notación. Decir que el intervalo (x_n, y_n), donde $x_n < y_n$, puede hacerse arbitrariamente pequeño en R_A significa que, cualquiera que sea ϵ (elemento positivo arbitrariamente próximo al elemento neutro 0), es posible determinar un entero N, función de ϵ, tal que $n > N(\epsilon)$ implique $y_n < x_n + \epsilon$ o $y_n - x_n < \epsilon$. La proximidad se traduce por una *diferencia* próxima al elemento neutro. En R_M^+ será necesario decir: cualquiera que sea η, superior al elemento neutro 1, pero arbitrariamente próximo a él, es posible determinar N en función de η, tal que

$$y_n < x_n \eta, \qquad \text{o} \qquad \frac{y_n}{x_n} < \eta.$$

En este caso es el *cociente* el que se halla próximo al elemento neutro 1.

Vamos a comprobar que B_2 se aplica en R_M^+ sirviéndonos del principio de los intervalos en R. Si $y_n < \eta x_n$, se tiene:

$$y_n - x_n < x_n(\eta - 1) < \alpha(\eta - 1)$$

ya que la sucesión x_n está acotada por uno cualquiera (α) de los y. Si n tiende a infinito, se ve que $y_n - x_n$ puede hacerse tan pequeño como se quiera, y lo mismo sucede con $\eta - 1$.

3. El isomorfismo fundamental.—Vamos a demostrar que R_A y R_M^+ son isomorfos, utilizando para ello el hecho de que ambos satisfacen a las condiciones A_1, A_2, B_1 y B_2.

Una observación sobre B_2.—Antes de iniciar esta demostración observemos que el grupo aditivo de los números racionales y el grupo multiplicativo de los racionales positivos satisfacen a A_1, A_2 y B_1 y que no son, sin embargo, isomorfos. En efecto, en este isomorfismo, + correspondería a $\times$, y cierto elemento b a cierto elemento c. El racional $b/2$ tal que $b/2+b/2=b$ debe corresponder a un elemento x tal que $x\cdot x=c$; pero, en general, no existe elemento racional x que satisfaga a $x^2=c$, lo que hace imposible el isomorfismo.

La propiedad fundamental que permite demostrar el teorema de isomorfismo es, pues, B_2; el principio de los intervalos es, en efecto, una propiedad esencial del cuerpo de los números reales.

Para establecer el isomorfismo partiremos de la «tabla»

$$\begin{array}{lccccccccc} R_A & \dots & -2 & -1 & 0 & 1 & 2 & \dots & n & \dots \\ R_M^+ & \dots & \dfrac{1}{a^2} & \dfrac{1}{a} & 1 & a & a^2 & \dots & a^n & \dots \end{array}$$

que se obtiene haciendo corresponder los elementos neutros (como en todo isomorfismo) y precisando que al número 1 de R_A le corresponda el a de R_M^+; a es arbitrario, pero lo supondremos mayor que 1, ya que el caso $a<1$ se deduce fácilmente (véase pág. 172). De las propiedades elementales de los grupos resulta que a a^p de R_M^+, siendo p entero relativo, le corresponde el entero p. Con mayor generalidad, la existencia de una raíz *enésima* en R nos lleva a asociar al número $\sqrt[n]{a^p}=a^{p/n}$ el número racional p/n. En la matemática elemental no se pasa de ahí y ello nos permite completar la ley de correspondencia entre los términos de una progresión aritmética en R_A con los de una progresión geométrica en R_M^+.

Con la ayuda del principio de los intervalos, establecido para los dos grupos, se puede asociar a todo $x\in R_A$ un $y\in R_M^+$. El orden arquimediano en R_A permite aproximar x en menos de $1/2^n$; es decir, definir los valores

$$\frac{p_n}{2^n}\leqslant x<\frac{p_n+1}{2^n},$$

donde p_n es entero. Pasando de n a $n+1$, se tiene, evidentemente,

$$\frac{p_n}{2^n}\leqslant\frac{p_{n+1}}{2^{n+1}}\leqslant x<\frac{p_{n+1}+1}{2^{n+1}}\leqslant\frac{p_n+1}{2^n}$$

y, puesto que la diferencia $\dfrac{p_n+1}{2^n}-\dfrac{p_n}{2^n}=\dfrac{1}{2^n}$ tiende a cero cuando n

tiende a infinito, tenemos una sucesión de intervalos encajados que definen x. Se asocia en R_M^+ una sucesión correspondiente de intervalos encajados considerando $a^{p_n/2^n}$, $a^{(p_n+1)/2^n}$. Son encajados en virtud de las propiedades de los exponentes fraccionarios y, según el lema del apartado precedente, el cociente de dos valores extremos, $a^{1/2^n}$, tiende a 1 cuando n tiende a infinito. Esta sucesión define un número y que, por definición, es el asociado de x.

La correspondencia obtenida de esta forma es evidentemente biunívoca, porque, de acuerdo con el orden arquimediano en R_M^+, todo elemento y de este grupo puede definirse por una sucesión de intervalos encajados $a^{p_n/2^n} \leqslant y < a^{(p_n+1)/2^n}$, de donde resulta una sucesión correspondiente en R_A.

Propiedades fundamentales. 1.° Demostremos que en esta correspondencia se conserva la ley de grupo. Si x está definido por

$$\frac{p_n}{2^n} \leqslant x < \frac{p_n+1}{2^n} \qquad \text{e } y \text{ por} \qquad a^{p_n/2^n} \leqslant y < a^{(p_n+1)/2^n};$$

x' por

$$\frac{p'_n}{2^n} \leqslant x' < \frac{p'_n+1}{2^n} \qquad \text{e } y' \text{ por} \qquad a^{p'_n/2^n} \leqslant y' < a^{(p'_n+1)/2^n},$$

se tiene:

$$\frac{p_n+p'_n}{2^n} \leqslant x+x' < \frac{p_n+p'_n+2}{2^n}$$

y

$$a^{(p_n+p'_n)/2^n} \leqslant yy' < a^{(p_n+p'_n+2)/2^n}.$$

Si n tiende a infinito, la longitud del intervalo que define $x+x'$ tiende hacia cero y se obtienen dos sucesiones correspondientes de intervalos encajados, una que define $x+x'$, y la otra, yy'.

La correspondencia $x \Longleftrightarrow y$ es, por tanto, un isomorfismo. En particular, al opuesto $-x$ de un número de R_A corresponde el inverso $1/y$ del transformado y de x.

2.° Si $a > 1$, la correspondencia conserva el orden en cada grupo. Esto quiere decir que, para $x' > x$, tenemos $y' > y$, o también, según A_1, que $x'-x > 0$ implica $y'/y > 1$. Pero los dos números $x'-x$ e y'/y se corresponden en el isomorfismo (según 1.°) y, siendo el primero positivo, la sucesión de intervalos

$$\left(\frac{p_n}{2^n}, \frac{p_n+1}{2^n} \right)$$

está formada por números positivos a partir de cierto lugar. La sucesión correspondiente en R_M^+ está formada por números mayores que 1. Se tiene, pues, $y'/y > 1$.

OBSERVACIÓN.—Supongamos que se establece un nuevo isomorfismo asociando al número 1 el número $a' = 1/a$. Sea y' el número asociado a $x \in R_A$, que está definido por la sucesión de intervalos $(a'^{p_n/2^n}, a'^{(p_n+1)/2^n})$ y tendremos $y' = 1/y$. Así, pues, y' es el número correspondiente a $-x$ en el primer isomorfismo. En este caso la correspondencia obtenida invierte el orden de los dos grupos.

4. Funciones logarítmica y exponencial.—El isomorfismo precedente permite definir dos funciones, según que tomemos como variable independiente el número x de R_A o el y de R_M^+. En el primer caso se obtiene la función exponencial de base a, denotada por $y = a^x$, y en el segundo, la función logarítmica de base a, simbolizada por $x = \log_a y$. De estas dos definiciones resulta que, para un número cualquiera de R_M^+, se tiene $y = a^{\log_a y}$.

Traduzcamos, completándolas, las propiedades del isomorfismo anterior, utilizando bien la notación exponencial, o bien la logarítmica.

Propiedades de la función exponencial.— 1.ª Como lo hemos señalado para $x = p/q$ racional, $y = a^{p/q}$ es igual a $\sqrt[q]{a^p}$, con lo que se encuentra de nuevo la notación de los exponentes fraccionarios. La función exponencial generaliza, pues, la notación exponencial y define a^x para x real cualquiera. Se tiene, además, $a^0 = 1$, $a^1 = a$.

2.ª Se verifica la propiedad fundamental $a^x \cdot a^{x'} = a^{x+x'}$, que se extiende por recurrencia al caso de n factores. En particular, $a^{-x} = \dfrac{1}{a^x}$.

3.ª Se deduce fácilmente que $(a^x)^{p/q} = a^{px/q}$, siendo p/q racional.

4.ª La función es creciente si $a > 1$, y decreciente si $a < 1$. En el primer caso tiende a infinito al tender x a infinito positivo, y a cero cuando x tiende a infinito negativo.

5.ª La función es continua; es decir, a todo número ϵ prefijado se puede hacer corresponder η tal que $|x - x_0| < \eta$ implique $|y - y_0| < \epsilon$. Basta utilizar la correspondencia inversa, $y \longrightarrow x$, observando que se conserva el orden. Si tenemos

$$y_0 - \epsilon < y < y_0 + \epsilon,$$

se deduce que

$$x_0 - \eta_1 < x < x_0 + \eta_2$$

y, eligiendo para η el menor de los números η_1 y η_2, la condición $|x - x_0| < \eta$ implica $|y - y_0| < \epsilon$.

6.ª Resulta de la continuidad que la propiedad 3.ª se extiende a un número real z cualquiera: $(a^x)^z = a^{xz}$. Es suficiente definir z por una sucesión regular de números racionales, $z_1, z_2, \ldots, z_n, \ldots$ Para cada uno de ellos, $(a^x)^{z_n} = a^{xz_n}$.

Si $z_n \longrightarrow z$, el primer miembro tiende hacia $(a^x)^z$ en virtud de la continuidad de la función exponencial de base a^x, y el segundo tiende a a^{xz} por la continuidad de la función exponencial de base a.

7.ª Se demuestra también que, para cualquier número real z, $(ab)^z = a^z \cdot b^z$.

Propiedades de la función logarítmica.—1.ª La función $x = \log_a y$ está solo definida para $y > 0$; $\log_a 1 = 0$, $\log_a a = 1$.

2.ª El isomorfismo se traduce por la propiedad fundamental

$$\log_a (y_1 y_2) = \log_a y_1 + \log_a y_2,$$

que se extiende al caso de n factores. En particular, $\log_a \frac{1}{y} = -\log_a y$ y, en consecuencia, $\log_a \frac{y_1}{y_2} = \log_a y_1 - \log_a y_2$.

3.ª Se deduce fácilmente que $\log_a y^{p/q} = p/q \log_a y$, cualquiera que sea el número racional p/q.

4.ª La función es creciente para $a > 1$. Tiende hacia $+\infty$ cuando y tiende a $+\infty$ y hacia $-\infty$ cuando y tiende a 0. Los resultados son inversos si es $a < 1$.

5.ª La función es continua; sigue siendo válida la demostración dada anteriormente intercambiando los papeles de x e y.

6.ª Resulta de la continuidad que la 3.ª propiedad conserva su validez para un número z cualquiera, $\log_a (y^z) = z \log_a y$. Es suficiente definir z por una sucesión regular de números racionales $z_1, z_2, \ldots, z_n, \ldots$, y se tiene $\log_a (y^{z_n}) = z_n \log_a y$. Si n tiende a infinito, y^{z_n} tiende hacia y^z (continuidad de la función exponencial) y el límite del primer miembro es $\log_a (y^z)$ (continuidad de la función logarítmica). El segundo miembro tiende hacia $z \log_a y$.

7.ª Resulta así la importante fórmula del cambio de base, que establece la relación entre $\log_a y$ y $\log_b y$, siendo a y b cualesquiera. Partiendo de la identidad $y = a^{\log_a y}$, se toman logaritmos en ambos miembros en el sistema de base b:

$$\log_b y = \log_a y \cdot \log_b a.$$

Se deduce así que los logaritmos en los dos sistemas son números proporcionales.

Señalemos que, de la fórmula precedente, resulta, para $y=b$,

$$\log_a b=\frac{1}{\log_b a}$$

EXTENSION DEL TEOREMA DE ISOMORFISMO

5. Caso de un grupo abstracto.—Hemos demostrado que R_A y R_M^+ son isomorfos utilizando el hecho de que ambos satisfacen las condiciones A_1, A_2, B_1 y B_2. Vamos a probar ahora, a costa de un esfuerzo suplementario justificado por la importancia del resultado, que todo grupo abeliano ordenado G es isomorfo a R_A si verifica las propiedades A_1, A_2, B_1, B_2, que repetiremos una vez más enunciándolas en su forma más abstracta.

A_1) *G es un grupo abeliano para una ley de composición denotada por* $*$. *Posee una relación de orden* $\leqslant$ *compatible con* $*$.

Dicho de otra manera, $x\leqslant y$ implica $x*z\leqslant y*z$ cualquiera que sea $z\in G$.

A_2) *Entre dos elementos de* G *existe una infinidad de otros.*

B_1) G *es arquimediano.*

Expresado de otra forma, si es e el elemento unidad del grupo, para $e<a$ la sucesión a, $a*a$, $a*a*a$, ... carece de mayorante, y termina por sobrepasar a cualquier elemento b a partir de cierto número n de factores a.

B_2) *Se aplica el principio de los intervalos*; es decir, si se tiene una doble infinidad de elementos x_i e y_i tales que

$$x_1<x_2<\ldots<x_{n-1}<x_n\ldots<y_n<y_{n-1}<\ldots<y_2<y_1$$

y el intervalo (x_n, y_n) puede hacerse tan pequeño como se quiera, existe entonces un elemento α, necesariamente único, común a todos los intervalos.

El hecho de que (x_n, y_n) sea arbitrariamente pequeño se traduce de la siguiente manera: Cualquiera que sea el elemento η superior al elemento neutro y tan próximo de él como queramos, se puede determinar $N(\eta)$ tal que $n>N(\eta)$ implique $y_n<x_n*\eta$ o bien $y_n*x_n'<\eta$, siendo x_n' el inverso de x_n.

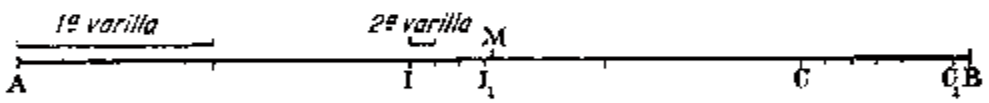
Esta fórmula contiene y generaliza las dos fórmulas precedentes de proximidad.

6. Consideraciones intuitivas.—Acabamos de ver que, si R_A y R_M^+ son efectivamente isomorfos, la división por 2 de un elemento de R_A corresponde a la extracción de la raíz cuadrada para su homólogo de R_M^+. En G nos llevará a resolver una ecuación del tipo $x * x = a$. Todo ello nos conduce a reflexionar sobre los conceptos de «media» y de «punto medio».

Deseamos marcar el punto medio de un trayecto rectilíneo AB y disponemos para ello de una serie de varillas de las que la primera es inferior a la mitad de AB, la segunda inferior a la mitad de la primera, y así sucesivamente.

Con la primera varilla jalonamos AB. Partimos de A y avanzamos hasta sobrepasar B. Si es C el último punto de número par $2n$ antes de B, marcamos con I el trazo de orden n: sabemos que I está a menos de una varilla por delante del punto medio M de AB. A continuación tomamos la segunda varilla y jalonamos a partir de C. Sea C_1 el último punto de número par $2n_1$ antes de B. Jalonamos a partir de I y marcamos el punto I_1, que lleva, después de I, el número n_1; sabemos que I_1 está a menos de una nueva varilla del punto medio M y delante de él.

Continuando con varillas cada vez más pequeñas obtendríamos una sucesión de puntos I_p, que tenderían hacia el punto medio buscado M. Introduciendo los puntos análogos J, J_1, etc., que sobrepasen el punto medio, vemos que dicho punto se determina por el principio de los intervalos.



En cierto modo, las demostraciones manejarán «varillas abstractas» Se observará la gran simplicidad técnica de estos razonamientos que se remontan a los primeros conceptos de *medida*.

El primer lema nos asegura que dada una varilla es siempre posible encontrar otra que sea menor que la mitad de la primera.

7. Propiedades del grupo G. LEMA 1.—*Para todo $x > e$ es posible hallar un elemento y tal que $e < y * y \leqslant x$.*

En efecto, según A_2, existen elementos entre e y x; sea z uno de ellos y formemos $z * z$. Si $z * z \leqslant x$, tomamos $y = z$. Si $x < z * z$ se considera el inverso z' de z y se toma $y = (x * z')$; se tiene así:

$$(x * z') * (x * z') = x * x * z' * z' < x * z' * z * z' * z = x.$$

Para una verdadera varilla x, esto significa romperla en dos trozos.

y ensayar la de longitud z. Si $z+z<x$ utilizamos z; en caso contrario, elegimos el otro trozo $x-z$. El papel de z' en la teoría general lo desempeña $-z$. Aconsejamos al lector que examine en concreto el caso en el que G es reemplazado por R_M^+.

LEMA 2.—*Si una sucesión indefinida de elementos x_n verifica la condición*

$$e < x_{n+1} * x_{n+1} \leqslant x_n$$

el elemento x_n tiende hacia el elemento neutro e.

Dicho de otra manera: cualquiera que sea el elemento b, arbitrariamente próximo a e, existen elementos de la sucesión x_n que penetran en (e, b). En efecto, si existiera un b inferior a todos los x_n, se tendría:

$$\begin{cases} b < x_2, \\ b * b < x_2 * x_2 \leqslant x_1; \end{cases}$$

del mismo modo, $b < x_3$,

$$b * b * b * b < x_3 * x_3 * x_3 * x_3 \leqslant x_2 * x_2 \leqslant x_1;$$

y, en general,

$$b * b * \ldots * b \text{ (con } 2^n \text{ factores)} < x_1,$$

lo que contradice el orden arquimediano.

LEMA 3.—*La ecuación $z * z = a$ admite en G una solución única.*

Esto equivale a la construcción del punto medio por el procedimiento de las varillas.

Supondremos $e < a$; el caso contrario se deduce pasando a los inversos.

El lema 1 nos asegura la posibilidad de encontrar un primer elemento u (primera varilla) tal que $e < u * u < a$; después, de manera general, una infinidad de otros u_n tales que

$$e < u_n * u_n \leqslant u_{n-1}.$$

En virtud del lema 1 se podrá formar la sucesión

$$u * u, \ \ldots, \ u * u * \ldots * u \quad (2n_1 \text{ elementos}),$$

que, según el orden arquimediano, acaba por sobrepasar a a. Si n_1 es el último entero tal que el producto de $2n_1$ elementos iguales a u no sobrepasa a a (piénsese en el punto C de la explicación), se elegirá

el elemento $u * u * \ldots * u$ (n_1 factores), que denotaremos por x_0. En estas condiciones, $x_0 * x_0 \leqslant a$ (imagínese el punto I de la explicación). Se toma igualmente el elemento $y_0 = u * u * \ldots * u * u$ ($n+1$ factores), y tendremos

$$x_0 * x_0 \leqslant a < y_0 * y_0.$$

Tenemos un primer intervalo (x_0, y_0) que encuadra al «punto medio» x desconocido.

Si la igualdad se cumple a la izquierda, la ecuación está resuelta. En otro caso, formamos la sucesión

$$x_0 * x_0, \qquad (x_0 * x_0) * (u_1 * u_1), \ldots, (x_0 * x_0) * (u_1 * u_1)^{2^p}, \ldots,$$

y se toma el último término que no sobrepase a a y el primero que la sobrepase. Designándolos, respectivamente, por x_1^2 e y_1^2, se tiene:

$$x_0 * x_0 \leqslant x_1 * x_1 < y_1 * y_1 < y_0 * y_0,$$

de donde,

$$x_0 \leqslant x_1 < y_1 < y_0 \qquad \text{con} \qquad y_1 = x_1 * u_1$$

y se proseguirá en idéntica forma. El problema quedará resuelto de manera sencilla si se realiza a la izquierda una igualdad. Si no, los intervalos encajados (x_n, y_n) definen, en virtud de B_2, un elemento z tal que $z * z = a$. Estos intervalos son, en efecto, arbitrariamente pequeños, ya que $y_n = x_n * u_n$ y u_n tiende hacia e, de acuerdo con el lema 2.

En cuanto a la unicidad de la solución x, resulta del hecho de que la función $x * x$ es estrictamente creciente sobre G.

8. Teorema de isomorfismo.—La posibilidad de resolver la ecuación $z * z = a$ en el grupo G conduce al isomorfismo entre R_A y G. La demostración es por completo análoga a la hecha para R_M^+. Haciendo $z = a^{1/2}$, se definirá la sucesión a, $a^{1/2}$, $a^{1/4}$, ..., $a^{1/2^n}$, ..., que tiende hacia e cuando n tiende a infinito, y se construye para todo $y \in G$ una sucesión de intervalos encajados $a^{p_n/2^n}$, $a^{(p_n+1)/2^n}$ que definen y. En R_A, la correspondiente sucesión

$$\left(\frac{p_n}{2^n}, \frac{p_n+1}{2^n}\right)$$

determina un elemento, x, que hacemos corresponder a y.

La correspondencia $x \Longleftrightarrow y$ define el isomorfismo buscado, porque

conserva la relación de orden (en la hipótesis $e < a$) y la ley de grupo en G como en R_A.

Si $a < e$, se ve fácilmente que $\leqslant$ en G corresponde a $\geqslant$ en R_A.

MEDIDA DE LONGITUDES SOBRE UNA RECTA

9. Extensión del teorema de isomorfismo.—El teorema se ha establecido entre un grupo abstracto, caracterizado por los axiomas A_1, A_2, B_1, B_2, y el grupo aditivo de los números reales. En ciertas cuestiones resulta cómodo considerar solo el conjunto R^+ de los números positivos, dotado de una ley de composición, la adición. Sobre este conjunto se verifican los axiomas siguientes (axiomas C).

C_1 *La ley de composición es asociativa, conmutativa y con elemento neutro* 0.

C_2 *Es un conjunto ordenado;* 0 *es el elemento más pequeño.*

C_3. *Para* $x < y$, *existe* z *tal que* $x + z = y$.

C_4. *El orden es compatible con la ley de composición:* $x < y$ *implica* $x + \alpha < y + \alpha$.

C_5. *Entre* 0 *y* $x > 0$, *y cualquiera que sea* x, *existe al menos un elemento.*

C_6. *El orden es arquimediano: si* $0 < x < y$, *existe un entero* n *tal que* $0 < y < nx$.

C_7. *Es aplicable el principio de los intervalos.*

Si de R^+ se pasa a un conjunto cualquiera que verifique estos axiomas, es fácil demostrar por un método enteramente análogo al precedente que entre estos dos conjuntos existe isomorfismo. El problema de la medida de una categoría de magnitudes constituye una aplicación importante del isomorfismo precedente. Examinemos el caso de las longitudes sobre una recta.

10. Propiedades de la recta.—No entra en nuestro propósito desarrollar una teoría axiomática de la recta. Sin embargo, para plantear adecuadamente el problema es útil recordar las propiedades de la recta estudiadas anteriormente, observando que se trata de nociones puramente geométricas cuya existencia es anterior a la medida de longitudes.

α) Sobre una recta dada existen tantos puntos como se quiera, y dos cualesquiera de ellos determinan la recta.

β) Entre los puntos de una recta existe una relación ternaria.

definida para todas las ternas posibles de puntos y que se expresa por la frase: «Tal punto está situado entre tal y tal otro.» Entre dos puntos A y B hay infinitos cuyo conjunto constituye el *segmento* AB. Se demuestra que la relación expresada por «el segmento NM no contiene al punto O» es una relación de equivalencia entre M y N, para un punto fijo O. Esta relación divide los puntos de la recta en dos clases: los pertenecientes a la clase M se dice que están *situados del mismo lado que* M *con relación a* O. Por último, cada clase se llama una semirrecta de origen O. Como se precisará algo más adelante, la relación ternaria permite definir sobre una semirrecta Ox una relación de orden por la cual O es anterior a todos los restantes puntos de aquella.

γ) Entre los segmentos de una recta existe una relación de equivalencia, que define la igualdad, y que, además de las tres propiedades habituales, satisface a las dos condiciones siguientes:

1.° $AB = BA$;

2.° si es C un punto cualquiera de la recta AB, sobre cada semirrecta que parte de C existe un punto D único tal que $AB = CD$.

δ) Dados tres puntos U, V, W, con V comprendido entre U y W, se dice que el segmento UW es la suma de UV y VW. Tal adición está definida solo para segmentos contiguos. Se demuestra (apoyándose en axiomas que no enunciaremos aquí) que permite asimismo definir para todos una adición entre las *clases de segmentos iguales;* la clase de AB viene representada por un segmento UV, y la de CD, por otro contiguo VW. Por abuso de lenguaje se hablará de adición de segmentos. Esta adición es asociativa, conmutativa y posee elemento neutro, constituído por la clase de los segmentos impropios del tipo AA, que resulta natural agregar a los segmentos propios del tipo AB.

Se ve que las propiedades α), β), γ), δ) implican que los segmentos OM de una semirrecta Ox verifican los cinco primeros axiomas C:

C_1 resulta de las propiedades enunciadas en δ);

C_2, porque se escribe $OA < OB$ si A está entre O y B;

C_3, porque para todo segmento $OA < OB$, se tiene que $OB = OA + AB$, y se tomará $OC = AB$;

C_4 resulta de C_3 y de la adición de las clases de segmentos iguales;

C_5, porque dado un segmento OA, siempre existe un punto B entre O y A.

Estas propiedades no permiten, sin embargo, definir el cociente de dos segmentos, ni, por consiguiente, el punto medio de un segmento. Para convencerse de ello, basta estudiar los contraejemplos

propuestos como ejercicios. Debido a esto, nos vemos obligados a introducir dos nuevos axiomas para la recta, que, por otra parte, son la formulación geométrica de C_6 y C_7.

Axioma de Arquímedes.—Si $OA < OB$, se construyen sobre la semirrecta los segmentos $OA = AA_1 = A_1A_2 = \ldots = A_{n-1}A_n = \ldots$, tales que A_i esté entre A_{i-1} y A_{i+1}. Se postula la existencia de un punto A_n tal que B esté entre O y A_n.

Axioma de Cantor-Dedekind.—Sea una recta cualquiera y elijamos sobre ella una sucesión indefinida de segmentos A_1B_1, A_2B_2, ..., etc., siendo el segmento $A_{n+1}B_{n+1}$ interior al A_nB_n, pudiendo hacerse A_nB_n menor que un segmento previamente dado, para n lo bastante grande.

En estas condiciones admitimos que existe un punto X interior a todos los segmentos A_nB_n.

La unicidad de X resulta inmediatamente de las hipótesis hechas en el enunciado: si existiera otro punto Y interior a todos los segmentos A_nB_n, siendo estos superiores a XY, no podrían hacerse arbitrariamente pequeños.

11. La medida de segmentos.—Queda, pues, asegurado el isomorfismo entre los segmentos OA y los números reales positivos. Resultará completamente determinado si se fija, en la aplicación biunívoca $OA \Longleftrightarrow a$, el segmento OA aplicado sobre el número 1; se tiene así la unidad de longitud. El número asociado a un segmento OA cualquiera (y a los segmentos congruentes con él) se llama su medida con la unidad de longitud elegida. Es también la relación entre el segmento considerado y el segmento unidad. Recíprocamente, existen siempre segmentos cuya medida es un número real positivo dado x.

Además, si el lector reconstruye la demostración del isomorfismo a partir de los axiomas C, verá que los lemas 1, 2 y 3 (págs. 175 y 176) adquieren sentido geométrico. El lema 1 demuestra que dado un segmento OA, existe OA_1 tal que

$$2\,OA_1 = OA_1 + OA_1 \leqslant OA.$$

El lema 2 prueba que, si $2\,OA_{n+1} \leqslant OA_n$, los segmentos de la sucesión OA, OA_1, ..., OA_n, ... pueden llegar a ser menores que cualquier segmento dado. Finalmente, el lema 3 prueba la existencia del *punto medio* de un segmento.

EJERCICIOS:

1. *Necesidad del axioma de Arquímedes.*—Llamemos punto A a un par ordenado (a_1, a_2) de números racionales, y recta, al conjunto de estos pares

ordenados lexicográficamente, definiendo la igualdad de segmentos AB y CD por las relaciones $|d_1-c_1|=|b_1-a_1|$, $|d_2-c_2|=|b_2-a_2|$.

Demuéstrese que, si O es (0, 0), A(1, 0) y B(0, 1), ninguno de los puntos A_n sobrepasa a B y no es posible medir OB tomando OA como unidad.

2. Cítese un contraejemplo que demuestre que si se admite el axioma de Arquímedes pero no el axioma de Cantor-Dedekind, no se puede afirmar que a todo número real corresponde un segmento que sea su medida respecto de una unidad dada.

EJERCICIOS SOBRE EL CAPITULO IX

9-1. Compruébese la fórmula

$$\log_2 7\cdot\log_3 7+\log_3 7\cdot\log_5 7+\log_2 7\cdot\log_5 7=\frac{\log_2 7\cdot\log_3 7\cdot\log_5 7}{\log_{30} 7}$$

9-2. ¿Cuál de los dos números, $\log_3 7$ y $\log_{\frac{1}{3}} 1/7$, es mayor?

9-3. Simplifíquese la expresión $b^{\frac{\log(\log b)}{\log b}}$, en la que los logaritmos están tomados en base a.

9-4. Resuélvase la ecuación $x^{\sqrt{x}}=(\sqrt{x})^x$.

9-5. Resuélvase la ecuación $\log_x 10+2\log_{10x} 10+3\log_{100x} 10=0$.

9-6. Resuélvase el sistema $7(\log_y x+\log_x y)=50$, $xy=256$.

9-7. Demuéstrese la fórmula

$$\log_a b\cdot\log_b c\cdot\log_c d\ \ldots\ \log_k l\cdot\log_l a=1.$$

9-8. Resuélvase la ecuación

$$4^x-3^{x-\frac{1}{2}}=3^{x+\frac{1}{2}}-2^{2x-1}.$$

9-9. Si los números $\log_k x$, $\log_m x$, $\log_n x$ están en progresión aritmética, pruébese que $n^2=(km)^{\log_k m}$.

9-10. Si los n números α, β, ..., λ son diferentes, demuéstrese que las n funciones $a^{\alpha x}$, $a^{\beta x}$, ..., $a^{\lambda x}$ son linealmente independientes.

9-11. Resuélvase la ecuación

$$\log_{2x}\frac{2}{x}\cdot(\log_2 x)^2+(\log_2 x)^4=1.$$

9-12. Resuélvase

$$\log_{\sqrt{2}} x\cdot\log_2 x\cdot\log_{2\sqrt{2}} x\cdot\log_4 x=54.$$

9-13. Resuélvase

$$\log_a x-\log_{a^2} x+\log_{a^4} x=\frac{3}{4}$$

9-14. ¿Cuál es la relación entre a, b, c que implica la igualdad

$$\log_{c+b} a+\log_{c-b} a=2\log_{c+b} a\cdot\log_{c-b} a?$$

9-15. Sea a un número real mayor que 1 y recuérdese que $\sqrt[n]{a}>1$ y que $\lim \sqrt[n]{a}=1$ cuando n tiende a infinito.

1.° Demuéstrese que la sucesión $u_n=n(\sqrt[n]{a}-1)$ es monótona decreciente y admite minorante. Hágase $t=\sqrt[n(n+1)]{a}$. Sea $\lambda(a)$ el límite de u_n cuando n tiende a infinito.

2.° Demuéstrese que si a y b son dos números tales que $a>b>1$,

$$\frac{a-b}{n\sqrt[n]{a^{n-1}}}<\sqrt[n]{a}-\sqrt[n]{b}<\frac{a-b}{n\sqrt[n]{b^{n-1}}}$$

3.° Acótese $\lambda(a+h)-\lambda(a)$; pruébese que $\lambda(a)$ es continua para $a>1$ y que $\dfrac{\lambda(a+h)-\lambda(a)}{h}$ tiene límite cuando h tiende a cero.

4.° Estúdiese el límite cuando n tiende a infinito de

$$w_n=n(\sqrt[n]{a}-1)(\sqrt[n]{b}-1)$$

y dedúzcase que $\lambda(ab)=\lambda(a)+\lambda(b)$.

9-16. *Conjunto triádico de Cantor.*—Sea un segmento AB de longitud unidad, que se divide en tres partes iguales AA_1, A_1B_1, B_1B y separamos la parte central, comprendidos los extremos. Igualmente, quitamos de AA_1 y B_1B sus partes centrales, etc. Así se prosigue indefinidamente.

Caracterícense las abscisas de los puntos que quedan, utilizando para ello el sistema de numeración de base 3. (*Resp.:* $0,\alpha_1\alpha_2\ldots\alpha_i\ldots$, con $\alpha_i=0$ o 2 pero no 1.)

9-17. *Conjunto perfecto de medida nula* (Borel).—Sea OA un segmento de 1 metro de longitud. Llamamos puntos decimétricos a los puntos de abscisas 0, 1, 2, ..., 10 dm, de los que hay once. Existen igualmente 101 puntos centimétricos, de los que 90 son nuevos, 1001 puntos milimétricos, de los cuales 900 son nuevos, etc.

Demuéstrese que es posible quitar un pequeño segmento alrededor de cada punto decimétrico, a continuación un segmento más pequeño alrededor de cada punto centimétrico, etc., de tal manera que quede *algo* del metro. Pruébese que la parte quitada puede hacerse tan pequeña como se quiera. (Los puntos 10^{-n}-métricos forman un conjunto de medida nula.)

9-18. *Lema de Bolzano-Weierstrass.*—Se dice que A es un *punto de acumulación* de un conjunto lineal si todo segmento que contiene a A contiene al menos un punto del conjunto distinto de A.

Demuéstrese que todo conjunto infinito de puntos de un segmento posee al menos un punto de acumulación (divídase en 2, 4, etc. partes y aplíquese el principio de los segmentos encajados).

9-19. *Aplicación del precedente.*—Sea a un número irracional positivo, x un entero positivo cualquiera, e y el mayor entero positivo contenido en ax.

1.° Demuéstrese que el conjunto de los $ax-y$ admite al menos un punto de acumulación.

2.° Dedúzcase sucesivamente que 0, y después que todo punto ax_0-y_0 del conjunto es punto de acumulación.

3.º Igualmente, que todo punto del segmento [0, 1] es también punto de acumulación.

4.º Conclúyase que una función continua $f(X)$ no puede tener dos períodos de cociente irracional (siendo p y q estos períodos y si los valores X_0 y X_1 no dan a la función valores iguales, se demostrará la existencia de dos enteros N y N' tales que $Np-N'q$ sea arbitrariamente próximo de x_1-x_0, de donde resulta una contradicción).

CAPITULO X

EL CUERPO DE LOS COMPLEJOS

CONSTRUCCION DEL CUERPO C

El conjunto R^2 de los pares ordenados (a, b) de números reales es un espacio vectorial sobre R. Puede definirse en R^2 una nueva operación, llamada producto, que asociada a la adición, dota a este conjunto de una estructura de cuerpo conmutativo. Se trata del *cuerpo de los complejos*, simbolizado por C.

1. Definición del producto.—El espacio R^2 es de dos dimensiones y los elementos $(1, 0)$ y $(0, 1)$ constituyen una base del mismo. Para simplificar, escribiremos:

$$\mathbf{e} = (1, 0), \qquad \mathbf{i} = (0, 1),$$

de donde,

$$(a, b) = a\mathbf{e} + b\mathbf{i}. \qquad [1]$$

La igualdad $(a, b) = (a', b')$ exige que $a = a'$ y $b = b'$. La suma $(a, b) + (a', b')$ se efectúa siguiendo el método habitual:

$$(a, b) + (a', b') = a\mathbf{e} + b\mathbf{i} + a'\mathbf{e} + b'\mathbf{i} = (a + a')\mathbf{e} + (b + b')\mathbf{i}.$$

La nueva operación que vamos a definir sobre R^2 se denomina *producto*, e impondremos *a priori* que verifique las dos propiedades siguientes:

α) ser distributiva con relación a la adición;
β) ser compatible con la ley externa definida sobre R^2; es decir,

$$[\lambda(a, b)]\,[\mu(c, d)] = \lambda\mu[(a, b)\,(c, d)],$$

donde

$$\lambda \in R \quad \text{y} \quad \mu \in R.$$

Según la fórmula [1] estas dos condiciones nos conducen a defi-

nir los productos entre sí de **e** por **i**. Lo hacemos de acuerdo con la tabla que sigue:

	e	**i**
e	**e**	**i**
i	**i**	$-\mathbf{e}$

Esta elección se explica por dos consideraciones: **e** es el elemento neutro identificado más adelante con 1, e $\mathbf{ii} = -\mathbf{e}$ corresponde a la introducción de un elemento, $\sqrt{-1}$, cuyo empleo se ha revelado coherente desde los primeros cálculos con los números imaginarios.

La ley así establecida para **e** e **i** es conmutativa. También es asociativa, puesto que [1]

$$\mathbf{e}(\mathbf{ie}) = (\mathbf{ie})\mathbf{e} = (\mathbf{ei})\mathbf{e}, \quad \mathbf{i}(\mathbf{ii}) = -\mathbf{i} = (\mathbf{ii})\mathbf{i}, \quad \mathbf{i}(\mathbf{ei}) = (\mathbf{ei})\mathbf{i} = (\mathbf{ie})\mathbf{i}, \text{ etc.}$$

Caso general.—Las condiciones $\alpha)$ y $\beta)$ permiten extender la definición de producto al caso de dos elementos cualesquiera:

$$\begin{aligned}(a\mathbf{e} + b\mathbf{i})(a'\mathbf{e} + b'\mathbf{i}) &= aa'\mathbf{e}^2 + ab'\mathbf{ei} + ba'\mathbf{ie} + bb'\mathbf{i}^2 \\ &= (aa' - bb')\mathbf{e} + (ab' + ba')\mathbf{i}, \end{aligned} \qquad [2]$$

o también

$$(a, b)(a', b') = (aa' - bb', ab' + ba').$$

Recíprocamente, la lineabilidad de la ley precedente con relación a los elementos a y b de una parte, y a los a' y b' de otra, demuestra que las condiciones $\alpha)$ y $\beta)$ se verifican efectivamente.

2. Propiedades del producto. Conmutatividad.— Asegurada para **e** y para **i** se generaliza inmediatamente al caso de dos elementos cualesquiera.

Asociatividad.—Al efectuar los dos productos

$$[(a\mathbf{e} + b\mathbf{i})(a'\mathbf{e} + b'\mathbf{i})](a''\mathbf{e} + b''\mathbf{i})$$

y

$$(a\mathbf{e} + b\mathbf{i})[(a'\mathbf{e} + b'\mathbf{i})(a''\mathbf{e} + b''\mathbf{i})]$$

[1] Escribimos sin signo el producto **(ei)** de dos vectores de R^2; no hay peligro de confusión con la ley externa $a\mathbf{e}$ o $b\mathbf{i}$ ya que los vectores se escriben en negrita.

nos vemos precisados, en uno y otro caso, a formar de todas las maneras posibles el producto de un elemento de cada factor. En cada uno de los términos así formados, el coeficiente ($\in$ R) es el mismo y la asociatividad para los elementos **e**, **i** implica la igualdad término a término de los desarrollos obtenidos, y, por ende, la de los dos productos.

Elemento neutro.—El elemento $\mathbf{e}=(1,0)$ es neutro para la multiplicación:

$$\mathbf{e}(a\mathbf{e}+b\mathbf{i})=a\mathbf{e}^2+b\mathbf{e}\mathbf{i}=a\mathbf{e}+b\mathbf{i}=(a\mathbf{e}+b\mathbf{i})\mathbf{e}.$$

Inverso.— Todo elemento distinto del elemento nulo (0, 0) admite un inverso para la ley multiplicativa.

Sea $a\mathbf{e}+b\mathbf{i}$; se trata de hallar $a'\mathbf{e}+b'\mathbf{i}$ tal que

$$(a\mathbf{e}+b\mathbf{i})(a'\mathbf{e}+b'\mathbf{i})=(aa'-bb')\mathbf{e}+(ab'+ba')\mathbf{i}=\mathbf{e}.$$

De donde resulta el sistema

$$aa'-bb'=1, \qquad ab'+ba'=0,$$

que admite la solución única:

$$a'=\frac{a}{a^2+b^2}, \qquad b'=\frac{-b}{a^2+b^2},$$

salvo que $a^2+b^2=0$, es decir (sobre R) $a=b=0$. El elemento nulo es el único que carece de inverso. Escribimos:

$$(a\mathbf{e}+b\mathbf{i})^{-1}=\frac{a}{a^2+b^2}\mathbf{e}-\frac{b}{a^2+b^2}\mathbf{i}.$$

En resumen, los elementos distintos de (0, 0) forman un grupo abeliano para la multiplicación.

3. Estructura de cuerpo.—El conjunto R^2, dotado de la adición y de la multiplicación, verifica todos los axiomas de la estructura de un cuerpo conmutativo. Es el cuerpo de los *números complejos*, denotado por C. Precisemos sus relaciones con R.

TEOREMA.—*El conjunto de los elementos de la forma* $(a, 0)$ *es un subcuerpo de* C, *isomorfo a* R.

Tenemos $(a, 0)=a\mathbf{e}$. Este subconjunto es cerrado respecto de la adición

$$a\mathbf{e}+a'\mathbf{e}=(a+a')\mathbf{e}$$

y también de la multiplicación

$$(a\mathbf{e})\,(a'\mathbf{e}) = aa'\mathbf{e}.$$

Por otra parte, la biyección $(a, 0) \Longleftrightarrow a$ hace de este subconjunto un subcuerpo isomorfo a R.

Convenio.—Los dos conjuntos anteriores difieren solo en sus notaciones, a en lugar de $(a, 0) = a\mathbf{e}$; convenimos por ello en identificarlos, lo que equivale a asimilar $\mathbf{e}$ y 1. El cuerpo C contiene entonces a R, que constituye uno de sus subcuerpos. Con este convenio, el elemento $\mathbf{i}^2 = -\mathbf{e}$ se asimila a -1. Sea entonces $a\mathbf{e} + b\mathbf{i}$ un número complejo cualquiera: $a\mathbf{e}$ se ha identificado con el número real a; en cuanto a $b\mathbf{i}$ representa hasta aquí el producto sobre $\mathbf{R}^2$ de $\mathbf{i} = (0, 1)$ por $b \in \mathbf{R}$. En virtud de las propiedades de la multiplicación sobre C, tenemos:

$$b\mathbf{i} = b(\mathbf{e}\mathbf{i}) = (b\mathbf{e})\mathbf{i}.$$

Identificando $b\mathbf{e}$ con b, según hemos convenido, asimilamos la operación externa de $\mathbf{R}^2$ a la multiplicación de los elementos b e $\mathbf{i}$ de C. Por razones de homogeneidad representaremos por i (en lugar de $\mathbf{i}$) el vector (0, 1), lo que nos conduce a la notación definitiva

$$z = a + bi.$$

Resumamos la teoría precedente.

TEOREMA.—*Salvo un isomorfismo, el cuerpo* R *puede sumergirse en un supercuerpo* C *que contiene un elemento* i *tal que* $i^2 = -1$. *Todo elemento de* C *se expresa de manera única en la forma* $a + bi$, *con* a *y* b *reales.*

Del teorema precedente resulta la posibilidad de aplicar las reglas del cálculo algebraico (cálculo sobre un cuerpo) al conjunto de los números reales completado con un elemento i cuyo cuadrado se sustituye por -1. Esto es lo que hicieron los creadores de los números imaginarios (véase después); pero el éxito de estos cálculos no debe hacer olvidar el verdadero carácter de los números complejos: son vectores de $\mathbf{R}^2$, como lo subraya bien la interpretación geométrica estudiada más adelante. Un abuso de notación, justificado por el isomorfismo precedente, aseguró el éxito de los algebristas italianos Cardano y Bombelli.

Historia.—Entiéndase bien que las reglas de cálculo sobre C no se han establecido *a priori:* su elaboración fue debida a una idea audaz de Cardano y

Bombelli, que, hacia 1550, utilizaron la cantidad «sofística» $\sqrt{-1}$ para resolver las ecuaciones de tercer grado. Euler introdujo la notación i hacia 1750. Es necesario llegar hasta el siglo XIX para hallar una justificación rigurosa de las reglas, con Gauss, al que se debe la palabra *complejo*, con Argand, que en 1806 creó la representación geométrica (véase pág. 193), con Cauchy, que da el método de las clases residuales (pág. 247), y Hamilton, la de los pares de números reales. El carácter misterioso que han tenido durante mucho tiempo los complejos lo atestigua su nombre de *imaginarios* que de hecho no han perdido todavía.

Notación.—Abandonemos la notación vectorial y representemos por una sola letra, con frecuencia z, el número $a+bi$. Sobre C, la igualdad $z=z'$ equivale a una doble igualdad sobre R, $a=a'$ y $b=b'$. El número a se llama parte real de $z(a=\Re z)$, y bi, su parte imaginaria ($bi=\Im z$). Un número de la forma bi es imaginario puro.

Observemos, finalmente, que C no es un cuerpo ordenado en el sentido del apartado 9 (pág. 108); el cuadrado del número i es igual a -1, que no puede pertenecer a la clase de los elementos positivos.

EJERCICIOS:

1. Determínense los números reales x e y tales que

$$(1-2i)x+(3+5i)y=1+3i.$$

2. Sobre C, resuélvase el sistema ($x, y, z \in$ C)

$$x+iy-2z=0, \qquad x-y+2iz=20, \qquad ix+3iy-(1+i)z=30.$$

3. Si es

$$j=-\frac{1}{2}+i\,\frac{\sqrt{3}}{2},$$

calcúlense j^2 y j^3. Dedúzcase de ello el valor de:

$$(a+bj+cj^2)(a+bj^2+cj); \qquad (a+b)(a+bj)(a+bj^2);$$
$$(aj^2+bj)(bj^2+aj); \qquad (a+bj+cj^2)^3+(a+bj^2+cj)^3$$

(a, b y c pueden ser complejos).

4. Determínense dos números complejos tales que uno de ellos tenga parte real dada, a, y el otro parte imaginaria también dada, bi, y cuyo producto sea igual al número $c+di$. Discusión.

4. Automorfismo de C: complejos conjugados.—Consideremos la aplicación de C en sí mismo que hace corresponder $\bar{z}=a-bi$ a $z=a+bi$. El número $\bar{z}$ se llama *complejo conjugado* del z. Se trata, evidentemente, de una biyección que coincide además con la biyección recíproca, ya que $\bar{\bar{z}}=z$. Los números reales son los únicos elementos inva-

riantes en esta aplicación. Se comprueba fácilmente que la suma y el producto de dos números complejos conjugados son reales:

$$z+\bar{z}=2a, \qquad z\bar{z}=a^2+b^2.$$

TEOREMA.—*La aplicación* $z \Longleftrightarrow \bar{z}$ *es un automorfismo del cuerpo* C. Esto quiere decir que en la aplicación $z \longrightarrow \bar{z}$ el conjugado de la suma es la suma de los conjugados, y el conjugado de un producto, el producto de los conjugados. Haciendo $z=a+ib$, $z'=a'+ib'$, se tiene:

$$\overline{(z+z')}=a+a'-i(b+b')=a-ib+a'-ib'=\bar{z}+\bar{z}'$$

e igualmente,

$$\overline{(zz')}=aa'-bb'-i(a'b+ab')$$

y

$$\bar{z}\bar{z}'=(a-ib)\,(a'-ib')=aa'-bb'-i(a'b+ab').$$

Se demuestra sin dificultad que el automorfismo precedente es el único existente en C, aparte del automorfismo neutro ($z \longrightarrow z$).

Consecuencias.—1.ª Las propiedades clásicas de un isomorfismo dan inmediatamente los transformados del opuesto y del inverso de un número z:

$$\overline{(-z)}=-\bar{z} \quad \text{y} \quad \overline{\left(\frac{1}{z}\right)}=\frac{1}{\bar{z}}$$

2.ª Si en una serie de operaciones sobre C (adiciones, multiplicaciones y sus inversas), todos los números se reemplazan por sus conjugados, los resultados son números complejos conjugados.

3.ª Las propiedades anteriores se utilizan en el cálculo de un cociente:

$$\frac{a+bi}{c+di}=\frac{(a+bi)\,(c-di)}{(c+di)\,(c-di)}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}\,i.$$

En particular, se tiene para el inverso, $\dfrac{1}{a+bi}=\dfrac{a-bi}{a^2+b^2}$

Módulo de un número complejo.—La introducción de los complejos conjugados permite dar una definición abstracta del módulo de un núme-

ro z, cuya interpretación se dará en el apartado 7. Sean $z=a+ib$ y $\bar{z}=a-ib$; el producto

$$z\bar{z}=a^2+b^2$$

es un número real positivo o nulo. La raíz cuadrada de esta suma se llama *módulo* o valor absoluto del número z y se representa por $|z|$; de donde,

$$|z|=\sqrt{z\bar{z}}=\sqrt{a^2+b^2}.$$

Observemos que se verifica siempre que

$$|a| \quad \text{o} \quad |b|\leqslant\sqrt{a^2+b^2} \quad \text{y también} \quad \mathcal{R}z\leqslant|z|.$$

Para un número real, el módulo coincide con su valor absoluto.

Propiedades.—α) La condición $|z|=0$ equivale a $z=0$.

β) El módulo de un producto es igual al producto de los módulos:

$$|zz'|=|z|\cdot|z'|.$$

En efecto,

$$|zz'|^2=zz'\cdot\overline{zz'}=zz'\cdot\bar{z}\cdot\bar{z'}=z\bar{z}\cdot z'\bar{z'}=|z|^2\cdot|z'|^2$$

y basta tomar la raíz cuadrada positiva.

γ) El módulo de una suma es inferior o igual a la suma de los módulos:

$$|z+z'|\leqslant|z|+|z'|.$$

En efecto,

$$|z+z'|^2=(z+z')(\bar{z}+\bar{z'})=z\bar{z}+z'\bar{z'}+z\bar{z'}+z'\bar{z}.$$

Los dos últimos términos del desarrollo son conjugados y, por consiguiente, su suma es real y vale $2\mathcal{R}(z\bar{z'})$; se sabe que $\mathcal{R}(z\bar{z'})\leqslant|z|\cdot|\bar{z'}|$ y, por tanto,

$$|z+z'|^2\leqslant|z|^2+|z'|^2+2|z|\cdot|z'|.$$

Basta extraer la raíz cuadrada.

La igualdad $|z+z'|=|z|+|z'|$ exige que $\mathcal{R}(z\bar{z'})=|z\bar{z'}|$, o

$$aa'+bb'=\sqrt{a^2+b^2}\sqrt{a'^2+b'^2},$$

es decir,

$$(ab'-ba')^2=0, \qquad \text{o sea} \qquad a'=\lambda a, \quad b'=\lambda b \quad \text{con} \quad \lambda>0.$$

Los dos números son, por tanto, proporcionales y su cociente es positivo.

EJERCICIOS:

1. Escríbanse en la forma $a+bi$ los números complejos siguientes:

$$(2+i)^7+(2-i)^7; \qquad (1+2i)^5-(1-2i)^5; \qquad \frac{1+i\operatorname{tg}\alpha}{1-i\operatorname{tg}\alpha}\ (\alpha \text{ real});$$

$$\frac{(1+2i)^2-(1-i)^3}{(3+2i)^3-(2+i)^2}; \qquad \frac{(1-i)^5-1}{(1+i)^5+1}; \qquad \frac{(1+i)^9}{(1-i)^7}.$$

2. Simplifíquense las expresiones:

$$(a+ib)^2+(a-ib)^2; \qquad (1+ai)^4+(1-ai)^4; \qquad \frac{a+ib}{c+id}+\frac{a-ib}{c-id}$$

3. ¿Existen números complejos fijos, a, b, c, d, tales que satisfagan la relación $\bar{z}=\dfrac{az+b}{cz+d}$ para todo $z \in C$?

4. Demuéstrese la afirmación del texto relativa a los automorfismos de C. Si es λ la imagen de i, se demostrará que $\lambda^2=-1$.

5. **Aplicaciones. Raíz cuadrada.**—Dado $z=a+bi$, todo número $u \in C$ tal que $u^2=z$ se llama raíz cuadrada de z. Poniendo $u=x+iy$, se tiene:

$$(x+iy)^2=a+ib.$$

Esta igualdad sobre C proporciona dos igualdades sobre R:

$$x^2-y^2=a, \qquad 2xy=b.$$

Elevando la segunda ecuación al cuadrado, tenemos:

$$x^2+(-y^2)=a, \qquad x^2(-y^2)=-\frac{b^2}{4},$$

con, además, signo de $xy=\epsilon_b$ ($+1$ si $b>0$ y -1 si $b<0$; $b=0$ queda provisionalmente excluido). El número x^2 es la raíz positiva, y el número $-y^2$, la raíz negativa de la ecuación

$$S^2-aS-\frac{b^2}{4}=0.$$

Se deduce de ella, respetando la condición suplementaria, que

$$u=x+iy=\pm\left[\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}}+i\epsilon_b\sqrt{\frac{\sqrt{a^2+b^2}-a}{2}}\right]$$

Se obtienen, pues, dos soluciones y solo dos; estas son números opuestos.

El caso excluido, $b=0$, exige $x=0$, o bien $y=0$. Para $a>0$ se tiene $u=x=\pm\sqrt{a}$ y, para a negativo, $u=\pm i\sqrt{-a}$. El caso $a=0$, $b=0$ da solo una sola raíz cuadrada, $u=0$, que cabe considerar como doble.

Se ve así que los números reales negativos admiten dos raíces cuadradas sobre el cuerpo C. Estas son números de la forma iy, con $y\in\mathrm{R}$, llamados con frecuencia *imaginarios puros*. En particular, el número -1 admite como raíces cuadradas $\pm i$, lo que conduce a la fórmula $i=\sqrt{-1}$ y a la antigua notación, $a+b\sqrt{-1}$, de los números *imaginarios*.

TEOREMA.—*Todo número complejo diferente de cero admite sobre* C *dos raíces cuadradas opuestas.*

OBSERVACIÓN.—La notación $u=\sqrt{z}$ representa en general una *cualquiera* de las raíces cuadradas del número z. El convenio habitual sobre R ($\sqrt{x}$ es la raíz positiva) pierde su significado sobre C, que no está ordenado. En algunas cuestiones, nos vemos obligados a elegir para $\sqrt{z}$ una u otra de las raíces cuadradas; convengamos en que $\sqrt{z}$ designe la raíz cuya parte real es positiva (signo + en la fórmula general).

Ecuación de segundo grado.—El método corriente de resolver la ecuación de segundo grado es válido sobre C. A partir de

$$az^2+bz+c=0, \qquad (a,\ b,\ c,\ \text{complejos})$$

resulta:

$$\left(z+\frac{b}{2a}\right)^2=\frac{b^2-4ac}{4a^2}$$

y el cálculo precedente demuestra que la ecuación es siempre posible aquí (la condición $b^2-4ac>0$ pierde todo significado sobre C). Se obtiene:

$$z=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

Si $b^2-4ac=0$, se dice que la ecuación tiene dos raíces confundidas.

TEOREMA.—*Sobre* C *una ecuación de segundo grado admite siempre dos raíces distintas o confundidas.*

Si z_1 y z_2 son estas raíces se verifica, igual que sobre R, que

$$az^2+bz+c=a(z-z_1)(z-z_2).$$

EJERCICIOS:

1. Calcúlense las raíces cuadradas de los siguientes números:

$$3-4i; \qquad -15+8i; \qquad -11+6i; \qquad 8+6i; \qquad 1-i\sqrt{3}.$$

2. Calcúlese $A=\sqrt{4+i}+\sqrt{4-i}$ adoptando el convenio precedente sobre el símbolo $\sqrt{\ }$. Indíquense todos los valores posibles de A si se deja indeterminada la elección de la raíz cuadrada.

3. Calcúlese $\sqrt[4]{2-i\sqrt{12}}$

[*Resp.:* 4 soluciones: $i^p\left(\frac{1+\sqrt{3}}{2}+\frac{1-\sqrt{3}}{2}i\right)$, con $p=0, 1, 2, 3$].

4. Resuélvanse las ecuaciones siguientes:

$$z^2-(2+i)z-1+7i=0; \qquad (2+i)z^2-(5-i)z+2-2i=0;$$
$$z^4+6z^3+9z^2+100=0.$$

REPRESENTACION TRIGONOMETRICA DE LOS NUMEROS COMPLEJOS

6. Imagen de un número complejo.—La geometría ordinaria del plano proporciona una imagen cómoda del espacio R^2. Elijamos un origen O y una base **u**, **v**, y hagamos corresponder al vector $(a, b) \in R^2$ el vector $\mathbf{OM}=a\mathbf{u}+b\mathbf{v}$ del plano. El grupo aditivo de R^2 resulta, de esta forma, isomorfo al grupo aditivo de los vectores del plano, o también al grupo de las traslaciones. Sin embargo, la multiplicación anteriormente definida solo encuentra una interpretación sencilla en el plano métrico. Así, la multiplicación por i,

$$i(a+ib)=-b+ai$$

hace corresponder al vector (a, b) el vector $(-b, a)$ que, en geometría métrica, se deduce del primero por rotación del ángulo $+\frac{\pi}{2}$ alrededor del origen.

Precisamos la interpretación precedente suponiendo **u** y **v** unitarios y ortogonales: $(\mathbf{u}, \mathbf{v})=+\frac{\pi}{2}$; el extremo de **u** se simboliza por 1 y el de **v**, por i. El punto M (a, b), que corresponde a $z=a+ib$ es su *imagen*; recíprocamente, z es el *afijo* del punto M. Con frecuencia, el punto se designa con la misma letra que el afijo.

7. Módulo y argumento.—La imagen M de $z=a+ib$ puede referirse al sistema de coordenadas polares asociado al sistema cartesiano (O, u, v). Poniendo $OM=\rho$, número real positivo, y designando por θ una cualquiera de las determinaciones del ángulo (Ox, OM) (definido módulo 2π), sabemos que

$$a=\rho\cos\theta, \qquad b=\rho\,\text{sen}\,\theta$$

y recíprocamente,

$$\rho=\sqrt{a^2+b^2}, \qquad \cos\theta=\frac{a}{\sqrt{a^2+b^2}}, \qquad \text{sen}\,\theta=\frac{b}{\sqrt{a^2+b^2}}$$

Vocabulario: ρ es el módulo de z, idéntico al número $\sqrt{z\bar{z}}$ considerado anteriormente, y θ es su argumento; se escribe:

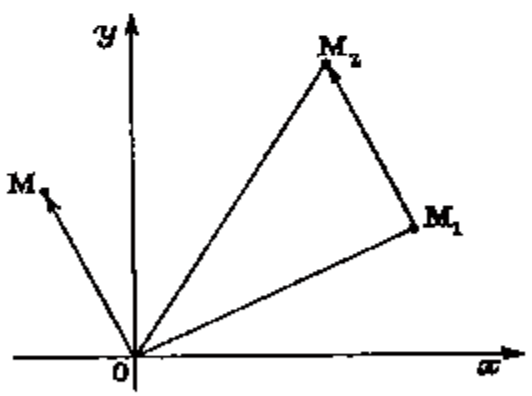
$$\rho=|z|, \qquad \theta=\arg z.$$

Con estas notaciones, z se escribe en su *forma trigonométrica*:

$$z=\rho\cos\theta+i\rho\,\text{sen}\,\theta=\rho(\cos\theta+i\,\text{sen}\,\theta)$$

y aparece como producto de un número positivo, su módulo, por otro número, $\cos\theta+i\,\text{sen}\,\theta$, de módulo 1 y de argumento θ.

El par de números (ρ, θ) puede interpretarse como los parámetros que definen una semejanza de punto doble O en el plano euclidiano: ρ es la razón de homotecia (aquí positiva siempre), y θ, el ángulo de rotación. Es la semejanza directa de punto doble O que hace pasar de la imagen del número 1 a la del número z. La correspondencia así definida es biunívoca.



OBSERVACIONES.—1.ª La interpretación geométrica del módulo permite encontrar de nuevo las propiedades α) y γ) (pág. 190). La condición $|z|=0$ exige que $OM=0$; por consiguiente, M confundido con O; el número $z=0$ tiene, por otra parte, un argumento indeterminado. La propiedad γ) se reduce a un teorema clásico de geometría.

2.ª Si z_1 y z_2 son los afijos de los puntos M_1 y M_2 la relación

$z = z_2 - z_1$ demuestra que la imagen de la diferencia $z_2 - z_1$ es el punto M tal que

$$\mathbf{OM} = \mathbf{OM_2} - \mathbf{OM_1} = \mathbf{M_1M_2}.$$

Prácticamente, el módulo y el argumento de $z_2 - z_1$ se leen sobre la figura sin que sea necesario construir M:

$$|z_2 - z_1| = M_1M_2, \qquad \arg(z_2 - z_1) = (\widehat{Ox,\ M_1M_2}).$$

EJERCICIOS:

1. Determínense el módulo y el argumento de cada uno de los números siguientes:

$$i; \qquad -1; \qquad 1+i; \qquad 1+i\sqrt{3}; \qquad -1+i\sqrt{3};$$
$$\sqrt{3}-i; \qquad 2+\sqrt{3}+i.$$

2. Precísese dónde se encuentran las imágenes de los números z tales que

$$|z| < 3, \qquad |z-i| \leqslant 1, \qquad |z-1+i| < 1.$$

3. Hállense los números complejos que verifican una u otra de las condiciones

$$|z| - z = 1 + 2i, \qquad |z| + z = 2 + i.$$

4. Todo número complejo z de módulo 1 puede escribirse en la forma $\dfrac{1+i\alpha}{1-i\alpha}$ donde α designa un número real. Calcúlese α, conociendo arg z.

5. Demuéstrese la fórmula

$$|z+z'|^2 + |z-z'|^2 = 2(|z|^2 + |z'|^2).$$

6. Lugar geométrico de la imagen de un número z que verifique a $|z| = 2|z-i|$.

7. Hállense los números z que satisfacen las dos relaciones

$$\left|\frac{z-12}{z-8i}\right| = \frac{5}{3}, \qquad \left|\frac{z-4}{z-8}\right| = 1. \qquad (\textit{Resp.: } z_1 = 6+8i,\ z_2 = 6+17i.)$$

8. Aplicación al grupo multiplicativo.—La notación anterior resulta particularmente cómoda para calcular un producto. Sea

$$z_1 = \rho_1(\cos\theta_1 + i \operatorname{sen}\theta_1), \qquad z_2 = \rho_2(\cos\theta_2 + i \operatorname{sen}\theta_2).$$

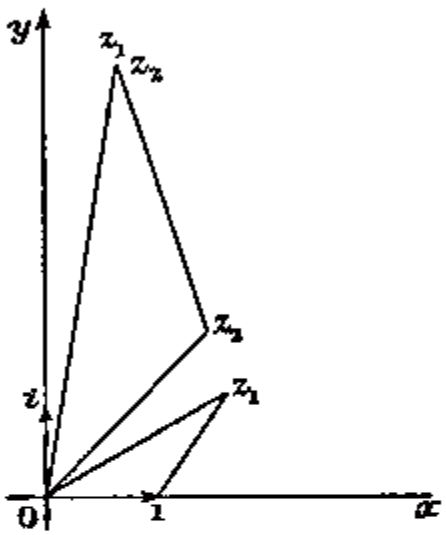
Se tiene:

$$z_1z_2 = (\rho_1\cos\theta_1 + i\rho_1 \operatorname{sen}\theta_1)(\rho_2\cos\theta_2 + i\rho_2 \operatorname{sen}\theta_2) =$$
$$= \rho_1\rho_2(\cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2) +$$
$$+ i\rho_1\rho_2(\cos\theta_1 \operatorname{sen}\theta_2 + \operatorname{sen}\theta_1\cos\theta_2).$$

En los paréntesis figuran desarrollos conocidos, que permiten escribir:

$$z_1z_2=\rho_1\rho_2\,[\cos(\theta_1+\theta_2)+i\,\operatorname{sen}(\theta_1+\theta_2)].$$

REGLA.—*El producto de dos números complejos tiene por módulo el producto de los módulos, y por argumento, la suma de los argumentos.*



Geométricamente, esto coincide con la ley de composición de semejanzas de centro fijo O, de donde el siguiente teorema.

TEOREMA.—*El grupo multiplicativo del cuerpo* C *es isomorfo al grupo de las semejanzas directas de punto doble dado* O *(imagen de* $z=0$*) en el plano euclidiano.*

Indiquemos algunas consecuencias.

α) La construcción de la imagen del producto z_1z_2 se deduce inmediatamente de la ley precedente: se aplica a la imagen de z_2 la semejanza definida por z_1 o, también, a la imagen de z_1 la semejanza definida por z_2.

β) La fórmula se extiende al producto de n números complejos:

$$z_j=\rho_j\,(\cos\theta_j+i\,\operatorname{sen}\theta_j)\qquad (j\text{ de }1\text{ a }n).$$

Tenemos:

$$z_1z_2\ldots z_n=\rho_1\rho_2\ldots\rho_n\,[\cos(\theta_1+\theta_2+\ldots+\theta_n)+i\,\operatorname{sen}(\theta_1+\theta_2+\ldots+\theta_n)].$$

Esta fórmula pone en evidencia las propiedades conmutativa y asociativa del producto.

γ) **Fórmula de De Moivre.**—Si los números $z_1, z_2, \ldots, z_n$ son iguales y de módulo 1, se obtiene la fórmula de De Moivre:

$$(\cos\theta+i\,\operatorname{sen}\theta)^n=\cos n\theta+i\,\operatorname{sen}n\theta.$$

δ) **Inverso y cociente.**—El isomorfismo precedente demuestra que el inverso del número de módulo ρ y argumento θ tiene por módulo $\dfrac{1}{\rho}$

y por argumento $-\theta$:

$$z^{-1}=\frac{1}{\rho}\left[\cos\left(-\theta\right)+i\,\mathrm{sen}\left(-\theta\right)\right].$$

En particular, el inverso de $\cos\theta+i\,\mathrm{sen}\,\theta$ es $\cos\theta-i\,\mathrm{sen}\,\theta$.

De lo anterior se deduce fácilmente el cociente de los números z_1 y z_2:

$$\frac{z_1}{z_2}=\frac{\rho_1}{\rho_2}\left[\cos\left(\theta_1-\theta_2\right)+i\,\mathrm{sen}\left(\theta_1-\theta_2\right)\right].$$

TEOREMA.—*El cociente de dos números complejos tiene por módulo el cociente de los módulos, y por argumento, la diferencia de los argumentos.*

e) La regla precedente permite extender la fórmula de De Moivre al caso de ser n negativo. Se tiene, en efecto,

$$\left(\cos\theta+i\,\mathrm{sen}\,\theta\right)^{-n}=\frac{1}{\left(\cos\theta+i\,\mathrm{sen}\,\theta\right)^{n}}=\frac{1}{\cos n\theta+i\,\mathrm{sen}\,n\theta}=$$
$$=\cos\left(-n\theta\right)+i\,\mathrm{sen}\left(-n\theta\right).$$

EJERCICIOS:

1. Demuéstrese que

$$\left(1+i\right)\left(1+i\sqrt{3}\right)\left(\cos\varphi+i\,\mathrm{sen}\,\varphi\right)=$$
$$=2\sqrt{2}\left[\cos\left(\frac{7\pi}{12}+\varphi\right)+i\,\mathrm{sen}\left(\frac{7\pi}{12}+\varphi\right)\right].$$

2. Calcúlense

$$\left(1+i\right)^{25},\qquad\left(\frac{1+i\sqrt{3}}{1-i}\right)^{10},\qquad\left(1-\frac{\sqrt{3}-i}{2}\right)^{12}.$$

3. Pruébese que

$$\left(1+i\right)^{n}=2^{n/2}\left(\cos\frac{n\pi}{4}+i\,\mathrm{sen}\,\frac{n\pi}{4}\right);$$

$$\left(\sqrt{3}-i\right)^{n}=2^{n}\left(\cos\frac{n\pi}{6}-i\,\mathrm{sen}\,\frac{n\pi}{6}\right).$$

4. Demuéstrese que, si $z+z^{-1}=2\cos\theta$, $z^{n}+z^{-n}=2\cos n\theta$

5. Simplifíquese la fracción $\left(\frac{1+i\,\mathrm{tg}\,\alpha}{1-i\,\mathrm{tg}\,\alpha}\right)^{n}$.

6. Calcúlense $\left(1+i\right)^{n}$, $\left(1+\cos\alpha+i\,\mathrm{sen}\,\alpha\right)^{n}$.

9. El plano de Argand-Cauchy.—Recapitulemos las propiedades de la representación geométrica de los números complejos. Un plano euclidiano referido a dos ejes coordenados rectangulares, graduados con la misma unidad, asocia a todo número $z = a + bi$ el punto M de coordenadas cartesianas (a, b) y de coordenadas polares (ρ, θ) $(\rho \geq 0)$. El plano de Cauchy resulta de una feliz confusión, dado que todo punto M, o todo número z, puede considerarse:

1.° Bien como un simple punto geométrico.

2.° Bien como un *operador de traslación;* la correspondiente al vector **OM** y que hace pasar de O a M.

Sumar z y z' se reduce a aplicar al punto geométrico M' la operación z-traslación, o también a aplicar a M la operación z'-traslación asociada a M'. Los resultados son idénticos.

3.° Bien como un *operador de semejanza*, que hace pasar del punto 1 [vector (0, 1)] al punto M.

Multiplicar z y z' es aplicar al punto M' la operación z-semejanza o también aplicar a M la operación z'-semejanza, lo que conduce al mismo resultado.

En particular, *i*-semejanza es una rotación de un cuarto de vuelta en sentido directo, de donde el interés que tiene en electricidad y electrotecnia para el estudio de las corrientes alternas

De este modo el cuerpo C posee su grupo aditivo isomorfo al grupo de las traslaciones del plano euclidiano y su grupo multiplicativo isomorfo al grupo de las semejanzas, los cuales pueden reunirse para constituir un cuerpo, ya que la multiplicación es distributiva con relación a la adición. Desde el punto de vista geométrico, ello se traduce por el hecho de que la adición geométrica de vectores es compatible con la semejanza: la transformada de la suma es la suma de las transformadas.

Lo que era válido para las traslaciones y semejanzas de centro O sobre un eje se extiende al plano: he aquí la razón por la que los complejos conservan todas las propiedades operativas de los números reales y pierden solo el orden.

EJERCICIOS:

1. ¿Por qué transformaciones geométricas se pasa de la imagen de z a la de z^{-1}?

2. Determínese el lugar geométrico del punto M, imagen de z, tal que las imágenes de i, z e iz estén alineadas.

3. Dados tres puntos A, B, C, de afijos respectivos a, b, c, demuéstrese

que una condición necesaria y suficiente para que el triángulo ABC sea equilátero y de sentido directo es que $a+bj+cj^2=0$.

4. ¿Cómo se desplazan las imágenes de las raíces de la ecuación $z^2-2\lambda z+1=0$ cuando λ real varía de $-\infty$ a $+\infty$?

5. Si es $z=\cos\varphi+i\,\text{sen}\,\varphi$, calcúlense el módulo y el argumento del número $u=z^2+z$. Lugar de su imagen cuando φ varía de 0 a 2π.

6. Sean M y M' las imágenes de z y z', ligados por la relación $zz'=a^2$, donde a designa un número real, y sean A y A' las imágenes de a y $-a$.

a) Indíquese la construcción de M' a partir de M; establézcase la relación, entre ángulos orientados de vectores,

$$(\widehat{\mathbf{MA'}},\ \widehat{\mathbf{MM'}})+(\widehat{\mathbf{MA}},\ \widehat{\mathbf{MO}})=2k\pi.$$

b) Calcúlense en función de x e y ($z=x+iy$) las componentes del vector $\mathbf{MM'}$, utilizando los números complejos.

7. Hállese el lugar de M, imagen de z, tal que z, z^2 y z^4 estén alineados.

10. Raíces *n*-ésimas de un número complejo.—Sea

$$z=\rho\,(\cos\theta+i\,\text{sen}\,\theta)$$

un número complejo dado. Todo número u tal que

$$u^n=z$$

se llama raíz n-ésima de z. La representación trigonométrica permite determinar los números u. Poniendo $u=r\,(\cos\alpha+i\,\text{sen}\,\alpha)$, se tiene:

$$r^n\,(\cos n\alpha+i\,\text{sen}\,n\alpha)=\rho\,(\cos\theta+i\,\text{sen}\,\theta),$$

de donde las dos igualdades, sobre R,

$$r^n=\rho, \qquad n\alpha=\theta+2k\pi,$$

designando por θ una de las determinaciones de arg z. De la primera igualdad se deduce:

$$r=\sqrt[n]{\rho},$$

porque todo número positivo admite una raíz positiva de orden n sobre R. Y, de la segunda, resulta finalmente:

$$\alpha=\frac{\theta}{n}+\frac{2k\pi}{n}$$

Dos valores de α congruentes módulo 2π determinan el mismo número u; se ve fácilmente que se obtienen n raíces n-ésimas y solo

n. Corresponden a los valores de k (mód. n): 0, 1, 2, ..., $n-1$, y se escriben:

$$u=\sqrt[n]{\rho}\left(\cos\frac{\theta+2k\pi}{n}+i\,\text{sen}\,\frac{\theta+2k\pi}{n}\right) \quad (k \text{ de } 0 \text{ a } n-1).$$

Las imágenes de estos números están sobre la circunferencia de centro O y radio $\sqrt[n]{\rho}$. Si M_0 corresponde a $u_0(k=0)$, las otras imágenes se deducen de ella mediante rotaciones de ángulos $\frac{2k\pi}{n}$ y forman los vértices de un polígono regular de n lados.

TEOREMA.—*Todo número complejo distinto de cero admite n raíces complejas cuyas imágenes son los vértices de un polígono regular de n lados.*

Raíces n-ésimas de la unidad.—En particular, las raíces n-ésimas de la unidad son las soluciones de la ecuación $z^n=1$. Vienen dadas por la fórmula

$$z_k=\cos\frac{2k\pi}{n}+i\,\text{sen}\,\frac{2k\pi}{n} \quad (k \text{ de } 0 \text{ a } n-1).$$

El polígono precedente está inscrito en la circunferencia de radio 1, siendo uno de sus vértices el punto $z_0=1$.

TEOREMA.—*Las n raíces n-ésimas de la unidad constituyen un grupo multiplicativo isomorfo al grupo de las rotaciones que conservan un polígono regular de n lados.*

Si $z_i^n=1$ y $z_k^n=1$ son dos raíces n-ésimas, se tiene que $(z_iz_k)^n=1$ y z_iz_k es también raíz n-ésima. Finalmente,

$$(z_i^{-1})^n=(z_i^n)^{-1}=1;$$

por consiguiente, el inverso de z_i es también raíz de la unidad. Tenemos, pues, un grupo de orden n. El isomorfismo con el grupo de las rotaciones de un polígono regular de n lados se establece haciendo corresponder a la raíz z_k la rotación $\frac{2k\pi}{n}$ que conserva el polígono. El grupo de las z_i está engendrado por las n primeras potencias de $z_1=\cos\frac{2\pi}{n}+i\,\text{sen}\,\frac{2\pi}{n}$:

$$z_0=(z_1)^0, \qquad z_2=(z_1)^2, \ \ldots, \ z_{n-1}=(z_1)^{n-1}.$$

El grupo se dice que es cíclico. Igualmente, el grupo de las rotaciones viene engendrado por las potencias de la rotación $\frac{2\pi}{n}$.

Cualquier otra raíz z_k ($k \neq 0$ y 1) engendra un subgrupo cíclico, que coincide con el grupo total si k es primo con n.

En particular, para $n=3$, tenemos:

$$z^3-1=(z-1)(z^2+z+1)=(z-1)(z-j)(z-j^2),$$

con

$$j=\cos\frac{2\pi}{3}+i\,\text{sen}\,\frac{2\pi}{3},\quad j^2=j=\cos\cdot\frac{2\pi}{3}-i\,\text{sen}\,\frac{2\pi}{3}$$

Los números j y j^2 son las raíces cúbicas imaginarias de la unidad. El siguiente teorema se refiere al caso general.

TEOREMA.—*Las raíces n-ésimas de un número complejo cualquiera son los productos de una de ellas por las raíces n-ésimas de la unidad.*

Si es u_0 una raíz particular de orden n de z, y u, una cualquiera de las otras, se tendrá:

$$u^n=u_0^n=z,\qquad \text{de donde}\qquad \left(\frac{u}{u_0}\right)^n=1;$$

el cociente u/u_0 es una cualquiera de las raíces n-ésimas de la unidad, con lo que el teorema queda demostrado.

EJERCICIOS:

1. Calcúlense

$$\sqrt[3]{-i},\qquad \sqrt[4]{1+i},\qquad \sqrt[3]{\frac{1-i}{\sqrt{3}+i}}$$

2. Demuéstrese que, para $a \neq 0$, las ecuaciones $(z-a)^m=b$ y $z^p-1=0$ no pueden tener más de dos raíces comunes, cualesquiera que sean los enteros m y p.

3. Resuélvase en el dominio complejo el sistema

$$x^n+y^n=\frac{c}{xyz},\qquad y^n+z^n=\frac{a}{xyz},\qquad z^n+x^n=\frac{b}{xyz}$$

4. Resuélvase la ecuación

$$z^4-z^3+z^2-z+1=0$$

5. Si α es una raíz n-ésima de la unidad, distinta de 1, demuéstrese que

$$1+\alpha+\alpha^2+\ldots+\alpha^{n-1}=0.$$

6. Si α es una raíz séptima de la unidad, distinta de 1, pruébese que

$$\frac{\alpha}{1+\alpha^2}+\frac{\alpha^2}{1+\alpha^4}+\frac{\alpha^3}{1+\alpha^6}=-2.$$

7. Determínense los números z tales que $2z=z_1+z_2$, donde z_1 y z_2 son dos raíces cúbicas distintas de z.

NOTA SOBRE LOS CUATERNIOS

11. Definiciones.—Es natural que surja la idea de construir «números» que desempeñen, con relación a la geometría del espacio euclidiano de tres dimensiones, un papel más o menos análogo al que tienen los complejos respecto de la geometría plana.

Si bien la adición vectorial continúa siendo estable para las semejanzas en el espacio, la composición de rotaciones que conservan el origen no es conmutativa, por lo que resulta necesario abandonar la conmutatividad.

Estos nuevos elementos son los *cuaternios* de Hamilton[1].

Sea K un cuerpo base, que podrá ser Q o R o C (véase ejercicio 8). Sean, por otra parte, cuatro elementos base, e, i, j, k (en ciertos casos e puede identificarse con 1).

	e	i	j	k
e	e	i	j	k
i	i	$-e$	k	$-j$
j	j	$-k$	$-e$	i
k	k	j	$-i$	$-e$

Consideremos el espacio vectorial sobre K que tenga como base $\{e, i, j, k\}$. Un vector del mismo será $ae+bi+cj+dk$ (a, b, c y d pertenecientes a K), y recibe el nombre de *cuaternio*.

Vamos a definir a continuación una *multiplicación* entre estos vectores. Para ello, comencemos por definir una multiplicación entre los elementos de la base, de acuerdo con la tabla adjunta.

Se observará en primer lugar que e es un elemento neutro, que i, j y k son tales que su cuadrado vale $-e$ y que la multiplicación

[1] Genial e imaginativo matemático irlandés. Impulsó el desarrollo de la óptica y de la mecánica (1805-1865).

es *anticonmutativa*:

$$\begin{aligned} ij &= k, & ji &= -k, \\ jk &= i, & kj &= -i, \\ ki &= j, & ik &= -j \end{aligned}$$

y, finalmente, que los ocho elementos e, i, j, k, $-e$, $-i$, $-j$, $-k$ forman un grupo. (Se recomienda al lector que complete la tabla y compruebe lo dicho.) El producto, evidentemente no conmutativo, de dos vectores $q = ae + bi + cj + dk$ y $q' = a'e + b'i + c'j + d'k$ se hace multiplicando cada término del primero por cada uno de los del segundo (¡atención al orden!) y haciendo la suma de los productos parciales. El resultado es el siguiente:

$$\begin{aligned} qq' = (aa' - bb' - cc' - dd')e + (ab' + ba' + cd' - dc')i + \\ + (ac' + ca' + da' - ad')j + (ad' + da' + ba' - ab')k. \end{aligned}$$

Propiedades de los cuaternios sobre el cuerpo de los reales.—Demostremos que, si el cuerpo base es el cuerpo de los números reales, el anillo de los cuaternios es un cuerpo.

A todo cuaternio, $q = ae + bi + cj + dk$, asociamos el cuaternio *conjugado* $\bar{q} = ae - bi - cj - dk$. Se demuestra sin dificultad que

$$q\bar{q} = (a^2 + b^2 + c^2 + d^2)e.$$

El número real $a^2 + b^2 + c^2 + d^2$ se llama *norma* de q, y la denotaremos por $|q\bar{q}|$. Solo el cuaternio nulo tiene norma nula. Se ve entonces que

$$q^{-1} = \frac{1}{|q\bar{q}|}\,\bar{q}.$$

Todo elemento no nulo admite, por tanto, un inverso. Los cuaternios de coeficientes reales forman un cuerpo no conmutativo. Se demuestra (Frobenius, 1878) que todo otro dominio no conmutativo que contenga a los números reales no puede ser un cuerpo.

Comparación con el cálculo vectorial.—En el cuaternio $q = ae + bi + cj + dk$, ae se llama parte escalar, y $bi + cj + dk$, parte puramente vectorial. A todo cuaternio se puede, por tanto, asociar un par formado por el escalar a y el vector $\mathbf{v}$ de componentes (b, c, d). Supongamos los ejes rectangulares y graduados con la misma unidad. Para dos cuaternios puramente vectoriales el producto se reduce a

$$qq' = -(bb' + cc' + dd')e + (cd' - dc')i + (da' - ad')j + (ba' - ab')k.$$

Se ve que la parte escalar es, salvo el signo, el producto escalar $\mathbf{v}\cdot\mathbf{v}'$, y que la parte puramente vectorial corresponde al producto vectorial $\mathbf{v}\wedge\mathbf{v}'$.

Los cuaternios, que gozaron de gran favor en la segunda mitad del siglo pasado, han cedido su lugar al álgebra y al análisis vectorial en cuanto se refiere a las cuestiones geométricas. En álgebra y en aritmética se utilizan cuaternios cuyos coeficientes pertenecen a diferentes cuerpos o anillos.

Véanse diversos complementos bajo forma de ejercicios (pág. 208, ejercicio 10-7).

OBSERVACIÓN.—El procedimiento empleado permite construir, a partir de cualquier tabla cuadrada que defina una multiplicación asociativa entre los elementos base de un espacio vectorial, un nuevo dominio con estructura de anillo, que se llama *sistema hipercomplejo* o, preferiblemente, *álgebra*.

Cuando los elementos base constituyen un grupo respecto de la multiplicación, se dice que se ha formado *el anillo del grupo* o *el álgebra del grupo*.

EJERCICIOS:

1. Resuélvase $\alpha x=\beta$ y $x\alpha=\beta$ para $\alpha=i$, $\beta=e+j$, y después para $\alpha=2e+j$, $\beta=3e+k$.

2. Demuéstrese que existe una infinidad de cuaternios que verifican la ecuación $x^2+e=0$. ¿Por qué este resultado no contradice el teorema relativo al número de raíces de una ecuación?

3. Resuélvase $x^2=e+i+j$.

4. Demuéstrese que todo cuaternio satisface a una ecuación de segundo grado con coeficientes reales.

5. Pruébese que la norma de un producto es igual al producto de las normas.

6. Demuéstrese que, si cada uno de los dos enteros m y n es la suma de cuatro cuadrados, lo mismo sucede con su producto.

7. Si el cuaternio x verifica la ecuación $x^2+ax+b=0$ de coeficientes reales, todos los cuaternios qxq^{-1} la satisfacen.

8. Demuéstrese que los cuaternios de componentes a, b, c, d complejas no forman un cuerpo (¡divisores de cero!). Convendrá designar $\sqrt{-1}$ por un símbolo distinto del utilizado para el cuaternio i.

9. Si $\bar{q}$ es el conjugado de q, la correspondencia $q \Longleftrightarrow \bar{q}$ no es un automorfismo del cuerpo de los cuaternios. ¿Por qué? Las propiedades que tiene esta correspondencia hacen que se denomine *antiautomorfismo* (al igual que se dice antidesplazamiento, o anticonmutativo cuando $xy=-yx$).

EJERCICIOS SOBRE EL CAPITULO X

10-1. Hállense los lugares geométricos de las imágenes M' y M'' de las raíces z' y z'' de la ecuación $z^2-2pz+c^2=0$, cuando la imagen P del número p describe la elipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-1=0, \qquad \text{con} \qquad c^2=a^2-b^2.$$

10-2. Precísense las condiciones necesarias y suficientes para que las raíces de la ecuación $x^2+(a+ib)x+c+id=0$ tengan el mismo argumento, o el mismo módulo.

Resp.: En uno y otro caso se obtiene la misma igualdad,

$$d(a^2-b^2)=2abc,$$

acompañada de una desigualdad $A\geqslant 0$ en el primer caso, $A\leqslant 0$ en el último, con $A=a^2+b^2-4\sqrt{c^2+d^2}$.

10-3. *Transformaciones circulares.*

1.° Demuéstrese que las transformaciones

$$z \to f(z)=\frac{az+b}{a'z+b'} \qquad (ab'-ba' \neq 0)$$

forman un grupo.

2.° Descompóngase la transformación $f(z)$ escribiéndola en la forma:

$$f(z)=\frac{a}{a'}+\frac{(ba'-ab')/a'^2}{z+b'/a'}$$

Interprétense las transformaciones simples del tipo

$$f(z)=z+z_0, \qquad f(z)=az, \qquad f(z)=\frac{1}{z} \qquad \text{(inversión} \times \text{simetría)}.$$

3.° Demuéstrese que estas transformaciones (llamadas circulares) son productos de un número *par* de inversiones positivas, propias o impropias, y recíprocamente.

4.° Estúdiense los puntos dobles de una transformación f y los dos haces conjugados de circunferencias definidos por estos puntos dobles.

5.° Sean α y β los afijos de los dos puntos dobles, que suponemos distintos; la transformación puede escribirse en la forma

$$\frac{Z-\alpha}{Z-\beta}=K\,\frac{z-\alpha}{z-\beta}$$

Dedúzcase de ello una clasificación de las transformaciones.

6.° Lugar de los puntos tales que

$$\arg\left(\frac{Z-\alpha}{Z-\beta}\right)\equiv 0 \pmod{2\pi}.$$

Lugar de los puntos tales que

$$\text{mód.}\ \frac{Z-\alpha}{Z-\beta}=\lambda.$$

OBSERVACIÓN.—Se sabe que el origen corresponde a $\varrho=0$, con θ indeterminado. Tiene interés «cerrar» el plano de Cauchy por un punto analagmático en el infinito, correspondiente a $\varrho=\infty$ y θ indeterminado. Este plano, distinto del plano euclidiano (abierto) y del plano argueslano (cerrado por la recta del infinito), corresponde por proyección estereográfica a una esfera.

10-4. *Fórmula de Euler.*

1.° Calcúlese $\operatorname{sen} n\alpha$ utilizando la fórmula de De Moivre. Dedúzcase de ella la relación

$$\operatorname{sen} n\alpha=\operatorname{sen}^n\alpha(C_n^1\cot^{n-1}\alpha-C_n^3\cot^{n-3}\alpha+C_n^5\cot^{n-5}\alpha-\ldots).$$

2.° Se considera el caso de n impar: $n=2m+1$. Demuéstrese que las raíces de la ecuación

$$C_{2m+1}^1x^m-C_{2m+1}^3x^{m-1}+C_{2m+1}^5x^{m-2}-\ldots=0$$

son

$$x_K=\cot^2\frac{K\pi}{2m+1}=\cot^2\frac{K\pi}{n},\qquad K=1,2,\ldots,m.$$

3.° Calcúlese la suma de las raíces de la ecuación. Haciendo uso de las relaciones

$$\frac{1}{\operatorname{sen}^2\alpha}=\cot^2\alpha+1\qquad\text{y}\qquad\operatorname{sen}\alpha<\alpha<\operatorname{tg}\alpha\ (\alpha\ \text{agudo}),$$

demuéstrese que

$$\frac{1}{6}(n-1)(n-2)<\left(\frac{n}{\pi}\right)^2+\left(\frac{n}{2\pi}\right)^2+\ldots+\left(\frac{n}{m\pi}\right)^2<\frac{1}{6}(n-1)(n+1).$$

4.° Dedúzcase el límite, cuando m tiende a infinito, de la sucesión

$$s_m=1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{m^2}$$

10-5. A.—Designamos por α y $\bar{\alpha}$ las raíces de la ecuación

$$z^2-2z+2=0.$$

1.° Calcúlense el módulo y el argumento de α^n y $\bar{\alpha}^n$.

2.° Escríbanse α^n y $\bar{\alpha}^n$ en la forma $a+bi$. (Se distinguirán los casos $n=4p$, $n=4p+1$, $n=4p+2$, $n=4p+3$.)

3.° Constrúyanse sus imágenes para $n\leqslant 8$.

B. 1.° Tomando $\alpha=1+i$, demuéstrese que todo número complejo z se expresa de manera única en la forma $z=x+y\alpha$, siendo x e y reales.

Demuéstrese la relación

$$z\alpha=-2y+(x+2y)\alpha.$$

2.° Se asocia a $z=x+y\alpha$ la matriz

$$M=\begin{bmatrix} x & y \\ -2y & x+2y \end{bmatrix}. \qquad [1]$$

Pruébese que el conjunto de las matrices M forma un cuerpo K', respecto de las operaciones de adición y multiplicación de matrices, y que este cuerpo K' es isomorfo al cuerpo K de los números complejos.

3.° Calcúlese M^n cuando $x=0$, $y=1$.

4.° ¿Constituye un cuerpo el conjunto de las matrices de la forma [1] si se suponen x e y complejos?

10-6. *Representación plana del grupo de las rotaciones en el espacio.* Sea $Oxyz$ un triedro trirrectángulo. La esfera

$$\Sigma=x^2+y^2+z^2-1=0$$

se somete a la proyección estereográfica S que tiene por polo el «polo sur» (0, 0, −1) y que transforma la esfera en su plano ecuatorial $z=0$. La imagen de $M(x, y, z)$ es $M'(x', y', 0)$.

1.° Exprésense las coordenadas de M en función de las de M'.

2.° Se introducen las cantidades $\zeta=x'+iy'$ e $\bar{\zeta}=x'-iy'$, llamadas coordenadas isótropas de M'. Demuéstrense las fórmulas

$$x+iy=\frac{2\zeta}{1+\zeta\bar{\zeta}}, \qquad x-iy=\frac{2\bar{\zeta}}{1+\zeta\bar{\zeta}}, \qquad z=\frac{1-\zeta\bar{\zeta}}{1+\zeta\bar{\zeta}}.$$

3.° Un eje Δ que pasa por O corta Σ en A y B. Una rotación R de ángulo φ alrededor de Δ transforma M en M_1. Estúdiese geométricamente la transformación $C=SRS$ que hace pasar de M' a M_1'. ¿Qué podemos decir de los puntos A' y B', de los haces de circunferencias que admiten estos como puntos fundamentales y como puntos límites? Dedúzcase de este estudio que C es una transformación circular del tipo

$$\zeta_1=\frac{\alpha\zeta+\beta}{\gamma\zeta+\delta},$$

definidos α, β, γ y δ salvo un factor. Demuéstrese que el grupo de las rotaciones R es isomorfo al grupo de las transformaciones de C.

4.° Se introducen coordenadas homogéneas haciendo $\zeta=\frac{\eta}{\xi}$, η y ξ definidas salvo un factor. Establézcanse las relaciones

$$\frac{x+iy}{2\eta\bar{\xi}}=\frac{x-iy}{2\xi\eta}=\frac{z}{\xi\bar{\xi}-\eta\bar{\eta}}=\frac{1}{\xi\bar{\xi}+\eta\bar{\eta}};$$

la transformación C se convierte en la transformación lineal

$$\begin{cases} \xi_1=\alpha\xi+\beta\eta, \\ \eta_1=\gamma\xi+\delta\eta. \end{cases}$$

Demuéstrese que a una rotación R corresponden *dos* matrices $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ normalizadas de modo que su determinante sea igual a $+1$. Dense las fórmulas explícitas para una rotación alrededor de Oz.

10-7. A.—Se considera el conjunto $\mathbb{Q}$ de pares $Q=(q_0;\ \mathbf{q})$ formados por un número (o *parte escalar*) q_0 y un vector libre (o *parte vectorial*) $\mathbf{q}$. Se definen en $\mathbb{Q}$:

a) Una adición tal que $Q+Q'=(q_0+q_0';\ \mathbf{q}+\mathbf{q}')$.

Identificando q_0 con el par $(q_0;\ 0)$ y q con $(0;\ q)$ escribiremos en adelante $Q=q_0+\mathbf{q}$. Justifíquese esta escritura.

b) Una multiplicación que denotamos por $*$ tal que, si $q_0=q_0'=0$,

$$\mathbf{q} * \mathbf{q}' = -\mathbf{q}\cdot\mathbf{q}' + \mathbf{q}\wedge\mathbf{q}' \quad [\cdot \text{ producto escalar, } \wedge \text{ producto vectorial}]$$

y dada, en general, por

$$Q * Q' = q_0q_0' + q_0\mathbf{q}' + q_0'\mathbf{q} + \mathbf{q} * \mathbf{q}'.$$

1. Demuéstrese que $\mathbb{Q}$ es un anillo no conmutativo.

2. Sea $\overline{Q}=q_0-\mathbf{q}$ el *conjugado* de $Q=q_0+\mathbf{q}$. Calcúlese el número $Q * \overline{Q}$, llamado *norma* de Q. Demuéstrese que $\mathbb{Q}$ es un cuerpo.

B.—Por un punto O pasa una recta Δ que orienta un vector unitario $\mathbf{u}$. Se considera la rotación $\mathcal{R}$ de eje Δ, de ángulo 2θ con relación a $\mathbf{u}$.

$\mathcal{R}$ cambia un punto M en otro M'.

1. Hállese por el estudio geométrico de la figura el valor reducido de

$$\mathbf{u}\wedge[(\mathbf{OM}+\mathbf{OM}')\operatorname{sen}\theta + (\mathbf{OM}-\mathbf{OM}')\cos\theta].$$

2. Compárese

$$(\cos\theta + \mathbf{u}\operatorname{sen}\theta) * \mathbf{OM} \quad \text{y} \quad \mathbf{OM}' * (\cos\theta + \mathbf{u}\operatorname{sen}\theta).$$

3. Se hace corresponder a $\mathcal{R}$ el par $Q=\cos\theta+\mathbf{u}\operatorname{sen}\theta$. Demuéstrese que

$$\mathbf{OM}' = Q * \mathbf{OM} * \overline{Q}$$

y hállese lo que de ello resulta en relación con el estudio y la composición de rotaciones en el espacio.

4. Utilizamos una referencia ortonormada: O punto de referencia, $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ vectores unitarios, y hacemos

$$\mathbf{E}=(1+0), \qquad \mathbf{I}=(0+\mathbf{i}), \qquad \mathbf{J}=(0+\mathbf{j}), \qquad \mathbf{K}=(0+\mathbf{k}).$$

¿Cuáles son las rotaciones que corresponden a estos pares base? Constrúyase la tabla de multiplicar para E, I, J, K con relación a $*$. Demuéstrese que permite el cálculo en $\mathbb{Q}$.

5. *Aplicación numérica:* siendo $\mathcal{R}$ la rotación de $+120°$ alrededor de Oy y $\mathcal{R}'$ la rotación de $+90°$ alrededor de $\mathbf{OA}=(6,\ 6,\ 7)$, determínense el eje y el ángulo de la rotación producto.

PARTE III

ALGEBRA DE LOS POLINOMIOS

CAPITULO XI

POLINOMIOS DE UNA VARIABLE

NOCIONES GENERALES

1. La noción clásica de polinomio.—En álgebra elemental, un polinomio es una función de un tipo determinado. El dominio de definición es un cuerpo conmutativo K, generalmente Q o R. Se define en primer lugar un monomio, expresión de la forma ax^n; es decir, en el lenguaje actual, la aplicación que a todo $x \in K$ hace corresponder ax^n, con $a \in K$ y $n \in N$. Un polinomio es la suma de cualquier número finito de monomios, y su valor para $x \in K$ puede escribirse siempre:

$$P(x) = a_0 + a_1 x + \ldots + a_n x^n \qquad (\text{con } a_n \neq 0). \qquad [1]$$

La definición de función nos ha llevado a distinguir entre f, la aplicación, y $f(x)$, valor de f para x. La teoría clásica mezcla estos dos aspectos. Si una aplicación de K en K, $x \longrightarrow f(x)$, *puede* revestir la forma [1], diremos que se trata de una *función racional entera*. Quede bien claro que toda suma de monomios define unívocamente una función de esa forma, pero el recíproco no es siempre cierto: una función racional entera determinada puede, si K es un cuerpo que encierra un número finito de elementos, expresarse de infinidad de maneras en forma de suma de monomios.

En la teoría que sigue abandonamos el aspecto función (que será tratado en la pág. 220) y tendremos en cuenta solo el aspecto *formal* del polinomio. En lógica estricta, deberemos abstenernos de considerar cualquier valor de la variable. Así despojado, el polinomio se convierte en un *polinomio formal*.

La teoría abstracta que se deduce permite, por otra parte, generalizar el estudio de las funciones racionales enteras, dejando indeterminada la elección del conjunto descrito por la variable.

2. Polinomio formal.—Los coeficientes de un polinomio pertenecen de ordinario a un cuerpo, a veces a un anillo [1] A, que suponemos con-

[1] El lector habrá manejado con frecuencia polinomios de coeficientes enteros; por tanto, elementos de Z.

mutativo. Con toda generalidad, adoptamos esta última hipótesis y justificaremos en el momento oportuno la restricción a un cuerpo.

Definición.—Se llama polinomio formal sobre un anillo A, toda sucesión indefinida de elementos de A en la que son nulos todos los términos a partir de cierto lugar. Los términos de esta sucesión se numeran comenzando por el índice 0, lo que implica un desfase de una unidad entre el orden de un término y el índice que le caracteriza. Se escribe, pues, en general:

$$P=(a_0,\ a_1,\ a_2,\ \ldots,\ a_n,\ 0,\ 0,\ \ldots)$$

denominando a_n el último término no nulo de la sucesión. Tenemos, por tanto, con certeza que $a_n \neq 0$, mientras que los términos que le preceden pueden ser nulos. La sucesión $O=(0, 0, \ldots, 0, \ldots)$ se considera también como un polinomio: el polinomio nulo. Los términos se llaman también coeficientes del polinomio. El número de términos diferentes de cero es finito, pero no está determinado el orden del último término no nulo. Cabe concebirlo tan elevado como se quiera, lo que justifica la escritura de una sucesión indefinida. De este modo, se puede siempre hablar de términos del mismo orden para dos polinomios cualesquiera. De manera abreviada, un polinomio P se representa por su término general de orden $i+1$; o sea, $P=[a_i]$. Designamos por $\mathfrak{P}$ el conjunto de los polinomios formales sobre el anillo A.

Grado.—Por definición, el índice n del último término no nulo de la sucesión se llama grado del polinomio. Todo polinomio tiene un grado determinado, con la única excepción del polinomio idénticamente nulo. La definición de grado carece de sentido para este polinomio.

Hablar del grado de un polinomio supone también implícitamente que este polinomio es diferente de O y exige con frecuencia algunas precisiones de vocabulario.

Igualdad.—Por definición, dos polinomios son iguales si, y solo si, los coeficientes del mismo orden son iguales. Esto implica la igualdad de los grados.

Monomio.—Recibe el nombre de monomio un polinomio en el que solo un coeficiente es distinto de cero. Un monomio queda, por consiguiente, caracterizado por su coeficiente y por su grado, que define el lugar en el que está colocado este coeficiente. A todo polinomio se

puede hacer corresponder cierto número de monomios, obtenidos tomando un coeficiente y uno solo no nulo del polinomio; p. ej.,

$$(a_0, 0, 0, \ldots), \quad (0, a_1, 0, \ldots), \quad \ldots, \quad (0, 0, \ldots, a_n, 0, \ldots).$$

Un polinomio de grado n contiene como máximo $n+1$ de ellos. Decimos que el polinomio ha sido «descompuesto» en monomios.

ESPACIO VECTORIAL DE LOS POLINOMIOS FORMALES SOBRE UN CUERPO

3. Adición.—Se define sobre $\mathfrak{P}$ una adición. Por definición, la suma de dos polinomios $P=[a_i]$ y $Q=[b_i]$ es el polinomio obtenido sumando los términos del mismo orden de P y Q:

$$P+Q=[a_i+b_i].$$

El número de términos no nulos de la sucesión así formada es ciertamente finito, y el resultado, un elemento de $\mathfrak{P}$. De esta forma el conjunto adquiere estructura de grupo conmutativo cuyo elemento neutro es el polinomio nulo $O=[0]$. El opuesto de P es $-P=[-a_i]$; la diferencia entre P y Q es

$$P-Q=P+(-Q).$$

TEOREMA.—*Todo polinomio es la suma de los monomios en los que se puede descomponer.*

Evidentemente, en virtud de la definición anterior,

$$P=(a_0, 0, 0, \ldots)+(0, a_1, 0, \ldots)+\ldots+(0, 0, \ldots, a_n, 0, \ldots),$$

en la que ciertos monomios pueden reducirse a O.

TEOREMA.—*El grado de la suma de dos polinomios, si existe, es inferior o igual al mayor de los grados de estos polinomios.*

En efecto, si los polinomios son de grados distintos, el grado de la suma es igual al mayor de los grados, ya que entonces la adición no modifica el último término (al menos) del polinomio que tiene mayor grado.

Si los polinomios son del mismo grado n, la suma es, en general, de grado n, salvo si $a_n+b_n=0$; en este caso será, como máximo, de grado $n-1$.

4. Multiplicación por un escalar.—Sea λ un elemento de A y $P=[a_i]$ un polinomio. Por definición, el polinomio λP, producto de P por λ, se deduce de P por multiplicación de todos sus coeficientes por λ:

$$\lambda P=[\lambda a_i].$$

Es evidente que se obtiene así un polinomio; pero, si A no es un anillo de integridad, el grado no se conserva necesariamente. Las tres propiedades siguientes:

$$(\lambda+\mu)P=\lambda P+\mu P, \qquad \lambda(P+Q)=\lambda P+\lambda Q,$$
$$\lambda(\mu P)=(\lambda\mu)P,$$

son consecuencias inmediatas de las reglas de cálculo sobre un anillo. Se dice que los elementos de A constituyen un dominio de operadores para los polinomios. Precisemos esto cuando A es un cuerpo.

5. Polinomios sobre un cuerpo K.—Si $A=K$, la operación externa antes definida verifica la propiedad suplementaria $1\cdot P=P$ y el conjunto $\mathfrak{P}$ dotado de la adición y de la ley anterior se convierte en un espacio vectorial sobre K.

Las nociones de dependencia e independencia (lineal) se aplican, por consiguiente, sobre $\mathfrak{P}$. Una relación de dependencia,

$$\lambda_1P_1+\lambda_2P_2+\ldots+\lambda_nP_n=O,$$

exige que el polinomio del primer miembro tenga todos sus coeficientes nulos. Si esta condición conduce a tomar los λ_α nulos, se dirá que los polinomios $P_1, P_2, \ldots, P_n$ son independientes.

Dimensión y base.—Es fácil demostrar que el espacio $\mathfrak{P}$ es de dimensión infinita: cualquiera que sea $n\in N$, siempre es posible hallar n polinomios independientes. Basta considerar la familia de polinomios cuyos coeficientes son todos nulos, a excepción de uno de ellos, que se elige igual a 1. Ponemos

$$P_0=(1,\ 0,\ 0,\ \ldots), \qquad P_1=(0,\ 1,\ 0,\ \ldots), \qquad P_2=(0,\ 0,\ 1,\ 0,\ \ldots)$$

y, en general, para P_i el coeficiente igual a 1 ocupa el lugar $(i+1)$ y P_i es de grado i.

Estos polinomios constituyen una familia libre, porque toda familia finita que se deduce de ellos es libre. En efecto, la igualdad

$\Sigma\lambda_i P_i = O$, en la que $0 \leqslant i \leqslant n$, exige $\lambda_i = 0$, ya que el polinomio P_i es el único para el que el coeficiente de orden $i+1$ es diferente de 0. Por otra parte, la familia de los polinomios P_i constituye una base de $\mathfrak{P}$, porque todo polinomio P es combinación lineal de un número finito de polinomios P_i:

$$P = \sum_{i=0}^{i=n} a_i P_i.$$

Subespacios.—Encontraremos numerosos subespacios de $\mathfrak{P}$, limitándonos ahora a observar que el conjunto de los polinomios de grado inferior o igual a un número dado n constituye un subespacio de $n+1$ dimensiones, evidentemente isomorfo con K^{n+1}.

NOTA.—Cuando A es un dominio de integridad (véase pág. 78) es posible definir los monomios P_0, P_1, ..., P_n, ..., ya que la multiplicación posee un elemento neutro. Todo polinomio $P=(a_0, a_1, \ldots, a_n, 0, \ldots)$ puede también escribirse:

$$P = \sum_{i}^{\cdot} a_i P_i,$$

pero el conjunto obtenido no es necesariamente un espacio vectorial.

ANILLO DE LOS POLINOMIOS SOBRE UN ANILLO

6. Producto de dos polinomios.—En el conjunto $\mathfrak{P}$ de los polinomios sobre un anillo A se define una nueva operación llamada *multiplicación* y denotada mediante los convenios habituales. Asociada a la adición, esta operación confiere a $\mathfrak{P}$ estructura de anillo.

Producto de dos monomios.— Por definición, el producto de dos monomios es un monomio cuyo grado es la suma de los grados, y el coeficiente, el producto, sobre A, de los coeficientes de los monomios considerados. Si uno de los monomios es nulo, también lo es el producto. Sean

$$p_i = (0, 0, \ldots, a_i, 0, \ldots) \quad \text{y} \quad p_j = (0, 0, \ldots, b_j, 0, \ldots);$$

se tiene:

$$p_{i+j} = p_i \cdot p_j = (0, 0, 0, \ldots, a_i b_j, 0, \ldots)$$

(el índice i, o el j, indica el grado del monomio considerado).

Observará el lector que la definición dada aquí *a priori* coincide con la regla clásica de multiplicación de monomios. Ha desaparecido la variable x, pero se ha reemplazado un cálculo por una definición.

Se demuestran inmediatamente las propiedades siguientes:

1.ª Si A es un anillo conmutativo, la multiplicación de monomios es conmutativa.

2.ª La multiplicación es asociativa:

$$(p_i \cdot p_j) \cdot p_k = p_i \cdot (p_j \cdot p_k).$$

Caso general.—Todo polinomio es una suma de monomios; por definición, el producto de dos polinomios P y Q se obtiene multiplicando cada monomio del primero, P, por cada uno de los del segundo, Q, y sumando los resultados obtenidos. Si

$$P = \sum_{i=0}^{i=n} p_i \quad \text{y} \quad Q = \sum_{j=0}^{j=p} q_j,$$

se tiene:

$$PQ = \sum_{i,j} p_i \cdot q_j.$$

La regla anterior equivale a considerar el conjunto producto de los monomios de P y de los monomios de Q, haciendo corresponder un monomio a todo elemento de este conjunto; a continuación se efectúa la suma de los monomios así obtenidos. El orden de los términos de la suma depende del método adoptado para formar el conjunto producto, pero esto no modifica el resultado.

En el producto PQ se obtiene el coeficiente c_k del monomio de grado k considerando todos los pares i y j tales que

$$i + j = k, \quad \text{con} \quad 0 \leqslant i \leqslant n, \quad 0 \leqslant j \leqslant p.$$

Si a_i y b_j designan coeficientes correspondientes de P y Q, se tendrá:

$$c_0 = a_0 b_0 \qquad c_1 = a_0 b_1 + a_1 b_0, \ \ldots, \ c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0,$$

donde se conviene que $a_k = 0$ (respectivamente, $b_k = 0$) si $k > n$ (respectivamente, $k > p$).

El último término se obtiene para $k = n + p$ y su coeficiente es $c_{n+p} = a_n b_p$. Si A no es un anillo de integridad, puede ocurrir que $c_{n+p} = 0$, incluso suponiendo $a_n \neq 0$ y $b_p \neq 0$.

TEOREMA.—*El producto de dos polinomios* P *y* Q *de grados* n *y* p *es un polinomio de grado* $n+p$ *como máximo, cuyo término de grado* k *admite como coeficiente* $c_k = \sum_{i+j=k} a_i b_j$, *siendo* a_i *y* b_j *los coeficientes de grados* i *y* j *de* P *y* Q.

7. Propiedades de la multiplicación. **1.ª Conmutatividad:** $PQ = QP$. Si en vez de efectuar PQ efectuamos QP, ello nos lleva a considerar el conjunto producto de los términos de Q y de los de P (escritos primero los monomios de Q), en lugar de a la inversa. Pero siendo conmutativa la multiplicación de monomios, los términos obtenidos para PQ y QP son los mismos, de donde resulta la conmutatividad en el caso general.

2.ª Asociatividad: $(PQ)R = P(QR)$. — Comparemos los dos productos (PQ)R y P(QR). En uno y otro producto deberemos formar el conjunto producto de los monomios de P, Q y R. Si (p_i, p_j, p_k) es uno de los elementos de este conjunto, para formar (PQ)R se efectúa primero el producto de los dos primeros monomios y se multiplica luego por el último. Al contrario, para formar P(QR) se multiplican primero los dos últimos monomios. En cualquier caso, la asociatividad del producto de monomios implica la de los polinomios.

3.ª Distributividad con relación a la adición: $P(Q+R) = PQ + PR$. — La igualdad de los coeficientes de los términos de grado k, en la fórmula siguiente, resulta inmediatamente de las reglas de cálculo sobre un anillo:

$$\sum_{i+j=k} a_i(b_j + c_j) = \sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i c_j.$$

(Bien entendido que ciertos coeficientes a_i, b_j, c_j pueden ser nulos según el lugar de k con relación a los grados de P, Q, R.)

8. Anillo de los polinomios.—Las propiedades precedentes hacen de $\mathfrak{P}$ un anillo conmutativo: el anillo de los polinomios sobre A.

TEOREMA. *Si* A *es un anillo de integridad, también el de los polinomios sobre* A *es un anillo de integridad.*

El producto de dos polinomios P y Q distintos de O, cuyos respectivos grados sean n y p, es en efecto un polinomio de grado $n+p$, ya que el último coeficiente $a_n b_p$ es con certeza diferente de cero en virtud de la hipótesis hecha respecto a A.

TEOREMA. — *Si* A *posee un elemento unidad e, el polinomio* $P_0=(e, 0, 0, \ldots)$ *es el elemento unidad en la multiplicación de polinomios.*

Esta propiedad se verifica para un monomio y, por consiguiente, también para un polinomio en virtud de la distributividad.

En resumen, si A es un dominio de integridad, también lo es el anillo de los polinomios sobre A. Precisaremos más adelante estas propiedades [1].

9. Anillo de los polinomios sobre un cuerpo.—Los polinomios P_1, P_2, ..., P_n, ..., definidos en la pág. 214, están ligados entre sí por la multiplicación. Se verifica, en efecto, que

$$P_2=(P_1)^2, \quad P_3=(P_1)^3, \ \ldots, \ P_n=(P_1)^n, \ \ldots$$

Por otra parte, siendo P_0 el elemento neutro, se puede poner $P_0=(P_1)^0$ (véase pág. 31). Por tanto, la base formada por los polinomios P_i está ahora confundida con el conjunto de las potencias positivas o nulas de P_1. Un polinomio cualquiera se escribe

$$P=a_0(P_1)^0+a_1(P_1)^1+a_2(P_1)^2+\ldots+a_n(P_1)^n.$$

Haciendo, para simplificar, $P_1=x$, encontramos la notación habitual

$$P=a_0x^0+a_1x^1+a_2x^2+\ldots+a_nx^n \qquad [1]$$

(salvo la presencia de x^0, que en la concepción clásica no se escribe, porque es igual a 1), pero con un significado diferente, ya que x no es aquí una variable, sino un polinomio particular.

Con todo, el anillo obtenido se llama *anillo de los polinomios de una variable sobre* K y se representa con la notación K[x].

El polinomio P, escrito en la forma [1], se dice ordenado con relación a las potencias crecientes de x. Este es el orden más natural, a pesar de que para ciertas operaciones (véase pág. 224) se prefiere el orden inverso. Con esta escritura, la multiplicación se efectúa como en álgebra elemental ($\lambda x^p \times \mu x^q=\lambda\mu x^{p+q}$) y es posible adoptar la disposición corriente.

OBSERVACIÓN.—K[x] es un dominio de integridad, pero no un cuerpo, ya que un elemento $P \neq O$ no tiene en general inverso. La relación

[1] Suponiendo que A sea un cuerpo, que es el caso más importante. Sin embargo, la notación $P=a_0x^0+\ldots+a_nx^n$ conserva su validez si A es un dominio de integridad.

$PQ = P_0$ exige, en efecto, que P y Q sean de grado cero; por consiguiente, $P = a_0 P_0$ y $Q = a_0^{-1} P_0$.

Además, el conjunto de los elementos de la forma $a_0 P_0$, con a_0 eventualmente nulo, constituye un subanillo K_0 de $K[x]$, isomorfo al cuerpo K, con el que a veces se identifica. Esto equivale a escribir de nuevo $x^0 = 1$, lo que se hace de manera natural si pensamos en x como elemento de K. De todas formas, la omisión de x^0 en un polinomio formal no puede originar confusión, ya que en los otros términos se pone de manifiesto su grado. Emplearemos, en general, esta escritura abreviada.

La relación de divisibilidad.—Las generalidades indicadas en la página 91 sobre la divisibilidad en los anillos de integridad con elemento unidad se aplican a $K[x]$. En particular, B divide a A (se escribe B|A) cuando existe un polinomio Q de $K[x]$ tal que

$$A = BQ.$$

Se dice también que A es un múltiplo de B. Los elementos unitarios son los elementos de K distintos de cero (o también los elementos λP_0). Dos polinomios asociados, A y B, son, por tanto, tales que $A = \lambda B$, donde $\lambda \in K$. En cada clase de elementos asociados se podrá, pues, elegir el representante que mejor se acomode a las necesidades prácticas. Una elección posible consiste en tomar el polinomio *normalizado*, es decir, tal que el coeficiente de la potencia más alta de x (término director) sea igual a 1. Señalemos que el producto de dos polinomios normalizados es normalizado.

Subanillos.—Se encuentran con frecuencia subanillos de $K[x]$. En particular, el conjunto de los múltiplos BQ de un polinomio dado B, con Q cualquiera de $K[x]$, constituye un subanillo; es cerrado respecto de la suma y del producto y contiene el opuesto de cada uno de sus elementos.

Casos particulares.—En la práctica, el lector encontrará principalmente los anillos $Q[x]$, $R[x]$, $C[x]$, que corresponden a los cuerpos Q, R, C. El primero, $Q[x]$, es un subanillo de $R[x]$, que, a su vez, es un subanillo de $C[x]$.

El automorfismo de C (véase pág. 188) se extiende por otra parte a $C[x]$. Se puede asociar a todo polinomio P de este anillo un *polinomio conjugado*, $\overline{P}$, cuyos coeficientes son los conjugados de los de P. La biyección $P \Longleftrightarrow \overline{P}$ es tal que

$$\overline{P+Q} \Longleftrightarrow \overline{P} + \overline{Q} \qquad \text{y} \qquad \overline{PQ} \Longleftrightarrow \overline{P}\,\overline{Q}.$$

Cualquiera que sea $P \in C[x]$, $(P+\overline{P})$ y $P\overline{P}$ pertenecen a $R[x]$.

10. Principio de especialización.—Sea $P = a_0 + a_1x + \ldots + a_nx^n$ un polinomio cuyos coeficientes pertenecen a un dominio de integridad A (en particular, un cuerpo K). Sea A' un superanillo de A, eventualmente confundido con A, o sea $A \subseteq A'$.

Si se hace corresponder al elemento x un elemento determinado, α, de A', el polinomio P define un nuevo elemento de A' llamado su valor para $x = \alpha$, y tal que

$$P(\alpha) = a_0 + a_1\alpha + \ldots + a_n\alpha^n.$$

Si α describe A', se obtiene así una aplicación de A' en A' llamada función racional entera. Todo polinomio define una función de este tipo. Indiquemos algunos ejemplos.

α) A es un cuerpo K, y A' se confunde con K; se vuelve a encontrar para $P(\alpha)$ el valor habitual y la definición clásica de polinomio.

β) A es un cuerpo K, y A' es el anillo $K[x]$ que, como hemos visto, contiene a K. Cualquier polinomio P define entonces una aplicación de $K[x]$ en $K[x]$, porque a todo $Q \in K[x]$ corresponde

$$P(Q) = a_0 + a_1Q + \ldots + a_nQ^n.$$

P(Q) se llama el compuesto de los polinomios P y Q. Se dice que se ha efectuado en P el cambio de variable $x = Q$.

La composición de polinomios así definida no es conmutativa, $P(Q) \neq Q(P)$. Para dos polinomios distintos de O se tiene:

$$\text{grado de } P(Q) = \text{grado de } P \text{ grado de } Q.$$

Finalmente, si $P(Q) = O$ y si grado de $Q \geqslant 1$, tendremos necesariamente $P = O$.

γ) A es un cuerpo K y A' el anillo de las matrices cuadradas de orden n sobre K (véase parte IV); el subconjunto de las matrices escalares aI es isomorfo a K; por consiguiente, para todo $x = M \in A'$, tenemos:

$$P(M) = a_0I + a_1M + \ldots + a_nM^n.$$

Se obtiene de esta forma una función racional de la matriz M.

En resumen, el mismo polinomio P conduce a definir funciones que pertenecen a diversos conjuntos, lo que justifica el estudio abstracto que aquí se hace. Además, la elección de las operaciones básicas se ha hecho de acuerdo con el cálculo sobre las funciones racionales en-

teras de una variable $\alpha \in A'$. Si P y Q son dos polinomios cualesquiera (con coeficientes en A), su suma, $S=P+Q$, y su producto, $T=PQ$, toman para $x=\alpha \in A'$ los valores

$$S(\alpha)=P(\alpha)+Q(\alpha), \qquad T(\alpha)=P(\alpha)\cdot Q(\alpha).$$

De lo anterior resulta la propiedad que sigue, llamada *principio de especialización*.

TEOREMA.—*Toda relación entre polinomios del anillo* A[x], *que haga intervenir la suma y el producto, también se verifica cuando reemplazamos el polinomio* x *por un elemento* α *de un anillo* A' *que contenga a* A.

EJERCICIOS:

1. Sobre el cuerpo de los enteros módulo 5 redúzcanse y ordénense los dos polinomios,

$$(2x^2+\bar{4}x+\bar{1})(3x^2+\bar{1}x+\bar{2}) \qquad y \qquad (-\bar{1}x^2+\bar{2}x+\bar{4})^2.$$

2. La misma cuestión con los siguientes polinomios, cuyos coeficientes pertenecen a Z/7:

$$(\bar{5}x^3+\bar{2}x-\bar{4})(\bar{6}x^2-\bar{3}x+\bar{1}); \qquad (\bar{3}x^2-\bar{4}x+\bar{2})^3.$$

3. ¿Cuántos polinomios distintos de segundo grado, de tercer grado, ..., de n-ésimo grado podemos formar sobre el cuerpo K=Z/3?

4. Si a, b, c, d pertenecen a un mismo cuerpo (de característica distinta de 2) establézcanse las condiciones necesarias y suficientes para que el polinomio

$$x^4+ax^3+bx^2+cx+d$$

sea un cuadrado perfecto.

$$\left(\textit{Resp.:}\ 2c=ab-\frac{a^3}{4}, \qquad c^2=a^2d. \right)$$

5. *Aplicación:* Sobre el cuerpo Z/7 determínense c y d de modo que

$$\bar{1}x^4+\bar{3}x^3+\bar{5}x^2+cx+d$$

sea un cuadrado perfecto; ¿de qué polinomio?

6. Sobre el cuerpo de los números racionales, demuéstrese que, mediante un cambio de variable, el polinomio

$$(8x+7)^2(4x+3)(x+1)$$

se reduce a un polinomio de segundo grado.

7. Dado un polinomio P(x) de tercer grado sobre un cuerpo K, si a, b, c designan tres elementos distintos de ese cuerpo, demuéstrese que

es posible determinar de manera única cuatro elementos, α, β, γ, δ, del mismo cuerpo, de modo que

$$P(x)=\alpha+\beta(x-a)+\gamma(x-a)(x-b)+\delta(x-a)(x-b)(x-c).$$

Generalícese.

8. Demuéstrese que, sobre el cuerpo Q, el polinomio

$$x^6+x^5+x^4+x^3-x^2+x-1$$

puede expresarse como producto de dos polinomios de tercer grado.

9. En el polinomio $ax^n+bx^{n-1}+\ldots+kx+l$ del anillo K[x] se efectúa el cambio de variable $x=y+h$, con $h\in K$. Demuéstrese que, en general, se puede elegir h de manera que el polinomio en y no admita término en y^{n-1}. Señálese el caso de excepción.

10. En el polinomio $nx^{n+2}-(n+2)x^{n+1}+(n+2)x-n$ se efectúa el cambio de variable $x=y+1$. Determínese el polinomio en y; dedúzcase de ello que el polinomio dado es divisible por $(x-1)^3$.

11. Calcúlese sobre C[x] el polinomio

$$P=(x+i)^n+(x+i)^{n-1}(x-i)+(x+i)^{n-2}(x-i)^2+\ldots+(x-i)^n.$$

12. Demuéstrese que, si K_1 es un subcuerpo de K, el anillo $K_1[x]$ es un subanillo de K[x].

CLASES DE CONGRUENCIA. DIVISION EUCLIDIANA

11. Clases de congruencia con relación a un polinomio.—Vamos a asociar a todo polinomio $B\neq O$ del anillo K[x] una relación de equivalencia análoga a la congruencia módulo n definida sobre Z.

Definición.— Dos polinomios A y A' se dicen congruentes módulo B si, y solo si, la diferencia $A-A'$ es un múltiplo de B:

$$A-A'=QB$$

y se escribe:

$$A\equiv A' \quad (\text{mód. } B).$$

La relación A congruente con A' (mód. B) obtenida de esta forma es evidentemente simétrica ($A'-A=-QB$), reflexiva ($A-A=OB$) y transitiva, porque

$$A-A'=QB$$

y

$$A'-A''=Q'B \Longrightarrow A-A''=(Q+Q')B.$$

Por consiguiente, es una relación de equivalencia que divide K[x] en clases disjuntas, llamadas clases de congruencia módulo B. Además, se ve inmediatamente cómo se forma una de dichas clases: la que contiene un polinomio determinado, A, está constituida por el conjunto de los polinomios $A+QB$, siendo Q un polinomio *arbitrario* de K[x]. Utilizaremos la indeterminación de Q para definir en cada clase un polinomio particular que la caracterice. Trataremos de hallar el polinomio de menor grado.

TEOREMA.—*En el anillo* K[x], *toda clase de congruencia módulo* B *contiene un polinomio y uno solo cuyo grado es inferior al de* B, *salvo la clase excepcional constituida por los múltiplos de B.*

Sea p el grado de B y supongamos que en una clase existe un polinomio R de grado inferior a p. Cualquier otro polinomio de esta clase se puede escribir en la forma $R+BQ$, donde BQ es de grado superior o igual a p (salvo para $Q=O$, que da R), y lo mismo sucede para cualquier otro polinomio de la clase considerada; de donde resulta la unicidad del polinomio R.

Busquemos ahora tal polinomio en una clase dada, $A+BQ$. Si es grado de $A<p$, bastará tomar $R=A$. Supondremos, pues, que es grado de $A\geqslant p$. Se procede por etapas sucesivas sin salir nunca de la clase considerada. Hallemos en primer lugar un polinomio A_1, de grado inferior en una unidad por lo menos al de A. Basta para ello determinar q_1 de manera que, en la diferencia

$$A_1=A-Bq_1,$$

desaparezca el término de mayor grado de A. Si a_nx^n es este término, siendo b^px_p el término análogo de B, la elección $q_1=\frac{a_n}{b_p}x^{n-p}$ cumple esta condición.

Si el grado de A_1 es inferior a p, tenemos $R=A_1$; en otro caso, se aplica a A_1 la misma transformación que a A, de donde resulta un polinomio $A_2=A_1-Bq_2$ de grado inferior al de A_1. Continuando así se pasa de A_2 a A_3, etc., y se construye una sucesión de polinomios A, A_1, A_2, A_3, ..., todos congruentes con A y de grados decrecientes. Cualesquiera que sean A y B, se presentará necesariamente una u otra de las circunstancias siguientes:

a) Uno de los polinomios de la sucesión es nulo: $A_h=O$. Es inmediato que todos los polinomios que le preceden son múltiplos de B y, en particular, $A=BQ_1$. La clase considerada es, por tanto, la de los

múltiplos de B; contiene al polinomio nulo, y todos los otros son de grado superior o igual a p.

b) Ninguno de los polinomios $A_1, A_2, \ldots, A_h$ es nulo. Puesto que los grados van decreciendo, después de $n-p+1$ operaciones, como máximo, se obtendrá un polinomio A_h de grado inferior a p. No es posible disminuir más el grado, y A_h es el polinomio R buscado. Queda, pues, demostrado el teorema.

12. División euclidiana.—Como el polinomio R, cuya existencia acabamos de demostrar, pertenece a la clase A, se podrá escribir:

$$A = BQ + R, \quad \text{con} \quad R = O, \quad \text{o bien grado de } R < \text{grado de } B.$$

Siendo único el polinomio R, el asociado Q está bien determinado. La obtención de los polinomios Q y R que verifican la igualdad anterior se llama *división euclidiana* de A (dividendo) por B (divisor). Además, Q resulta del cálculo anterior: sumando miembro a miembro las igualdades que definen $A_1, A_2, \ldots, A_h - R$, se obtiene:

$$R = A - B(q_1 + q_2 + \ldots + q_h),$$

de donde $Q = q_1 + q_2 + \ldots + q_h$.

Prácticamente, la operación se hace adoptando la disposición clásica de la división y ordenando los polinomios de acuerdo con las potencias decrecientes. Se determinan los términos sucesivos del cociente (de grados decrecientes) y se forma la sucesión de restos parciales $A_1, A_2, \ldots, A_h = R$. La operación termina cuando el grado de A_h es inferior a p.

EJEMPLO.—Sobre el cuerpo de los racionales, efectúese la división de $A = -x^3 + 2x^2 + 2x + 1$ por $B = -2x^2 + 2x - 1$. Se escribe:

$$\begin{array}{l|l} -x^3 + 2x^2 + 2x + 1 & -2x^2 + 2x - 1 \\ \hline -x^3 + x^2 - \frac{1}{2}x & +\frac{1}{2}x - \frac{1}{2} \\ \hline x^2 + \frac{5}{2}x + 1 & \\ x^2 - x + \frac{1}{2} & \\ \hline \frac{7}{2}x + \frac{1}{2} & \end{array}$$

de donde,

$$-x^3+2x^2+2x+1=(-2x^2+2x-1)\left(\frac{1}{2}x-\frac{1}{2}\right)+\frac{7}{2}x+\frac{1}{2}$$

OBSERVACIONES.—1.ª Los coeficientes del cociente y del resto se obtienen por una sucesión de operaciones definidas sobre un cuerpo: adición, multiplicación y sus inversas. De ello resulta que estos coeficientes pertenecen al menor cuerpo que contiene los coeficientes de los polinomios considerados, y ello independientemente del cuerpo base elegido para definir el conjunto K[x]. Así, p. ej., sobre el cuerpo de los números reales, el cociente y el resto de dos polinomios de coeficientes racionales son racionales.

2.ª En general, la operación no es posible cuando los coeficientes se eligen en un anillo. Sin embargo, si el coeficiente del término de mayor grado de B es un elemento unitario del anillo, se comprueba fácilmente que los coeficientes del cociente y del resto pertenecen a este anillo; p. ej., la división de dos polinomios con coeficientes enteros se hace sobre el anillo Z cuando el coeficiente del término de mayor grado de B es ± 1.

EJERCICIOS:

1. Efectúese la división de $A=x^4+6x^3+10x^2+3x-6$ por $B=x^2+3x$. Dedúzcase de ello la descomposición de A en dos trinomios de segundo grado.

2. Efectúese sobre C la división de $A=x^2-3ix-5(1+i)$ por $B=x-1+i$.

3. Determínense m y p de manera que x^3+px+q sea divisible por x^2+mx-1. Escríbase la igualdad que resulta.

4. Hágase la división de x^m-a^m por x^p-a^p y dedúzcase la condición para que el resto sea nulo.

5. Lo mismo para el polinomio x^m+a^m, dividido por x^p+a^p.

6. ¿Cuál es el valor numérico de $A=2x^4-4x^3-7x-14$ para $x=1+\sqrt{3}$? Divídase A por un polinomio B (de coeficientes racionales) que se anule para el mismo valor de x.

7. Sobre Q[x], efectúese la división del polinomio $nx^{n+1}-(n+1)x^n+1$ por $(x-1)^2$. Demuéstrese que el resto es nulo.

13. Propiedades.—Se obtienen propiedades análogas a las de la división de números enteros.

TEOREMA I.—*Si en una división se multiplica el dividendo y el divisor por un mismo polinomio, el cociente no cambia, pero el resto queda multiplicado por este polinomio.*

La igualdad $A = BQ + R$ implica $AC = (BC)Q + RC$ y, además, grado de $RC <$ grado de BC, o $RC = O$.

TEOREMA II.—*Todo polinomio* C *que divide a dos polinomios,* A *y* B, *divide también al resto de la división de* A *por* B.

Si $A = CA_1$ y $B = CB_1$, en la igualdad $A = BQ + R$ tendremos $R = C(A_1 - BQ_1)$.

TEOREMA III.—*En la división de la suma de dos polinomios por un tercero, el cociente es la suma de los cocientes, y el resto, la suma de los restos.*

TEOREMA IV.—*El resto de la división del producto de dos polinomios por un tercero es el mismo que el de la división del producto de los restos.*

Las igualdades

$$A = BQ + R \qquad \text{y} \qquad A' = BQ' + R'$$

implican

$$AA' = BQ'' + RR' \quad \text{o} \quad AA' \equiv RR' \text{ (mód. B)}.$$

Los polinomios AA' y RR' pertenecen, pues, a una misma clase de congruencia; pero, en general, el grado de RR' es superior al de B.

14. División por $x - a$.—Sea $A \in K[x]$; si se divide por el binomio $x - a$, con $a \in K$, el resto es generalmente de grado cero y, por tanto, un elemento λ de K. Se tiene:

$$A = (x - a)Q + \lambda.$$

Reemplazando x por $a \in K$, de acuerdo con el principio de especialización, resulta:

$$A(a) = 0 \cdot Q(a) + \lambda, \quad \text{de donde} \quad \lambda = A(a).$$

TEOREMA.—*El resto de la división de un polinomio por* $x - a$ *se obtiene reemplazando* x *por* a *en aquel.*

Consecuencia.—*Es condición necesaria y suficiente para que un polinomio sea divisible por* $x - a$ *que se anule cuando en él se sustituye* x *por* a.

El elemento a de K se llama *raíz* del polinomio.

TEOREMA.—*Sobre un cuerpo conmutativo* K, *un polinomio no nulo de grado n posee como máximo n raíces.*

El teorema es cierto para $n=1$. Suponiéndolo cierto para $n-1$, demostremos que lo es también para n. Si $x=a$ es una raíz de A [1], tenemos:

$$A=(x-a)Q$$

y Q es de grado $n-1$. Cuando reemplazamos x por un elemento cualquiera de K esta igualdad formal se convierte en una igualdad sobre K y, por consiguiente, el segundo miembro podrá anularse solo si se anula uno de los factores. Por tanto, las raíces de A son, además de a, las del polinomio Q, cuyo número es $n-1$ como máximo. El polinomio A tiene, pues, n raíces a lo sumo.

OBSERVACIÓN.—Sean P y $\bar{P}$ dos polinomios conjugados del anillo $C[x]$. Tomando para x valores complejos conjugados, z y $\bar{z}$, se demuestra sin dificultad que $P(z)$ y $\bar{P}(\bar{z})$ son complejos conjugados. Según el teorema anterior, resulta de ello que, si z es raíz de P, $\bar{z}$ lo es de $\bar{P}$.

Consecuencias.—Si a, b, ..., l son raíces distintas del polinomio A, este es divisible por el producto $B=(x-a)(x-b)\dots(x-l)$. En efecto: $A=(x-a)Q_1$ y, haciendo $x=b$, $A(b)=0$, y como en el segundo miembro $b-a\neq 0$, se tendrá $Q_1(b)=0$; por tanto, $Q_1=(x-b)Q_2$ y $A=(x-a)(x-b)Q_2$. Se continúa haciendo $x=c$, ..., etc.

15. Polinomios y funciones racionales enteras.—Se sabe que todo polinomio P de $K[x]$ define una función racional entera de $x\in K$. Esta es

$$x\longrightarrow P(x)=a_0+a_1x+\dots+a_nx^n.$$

La recíproca no es en general cierta: a dos polinomios distintos puede corresponder la misma función racional entera. Vamos a verlo en un ejemplo; siendo K el cuerpo de los enteros módulo 3, consideremos los dos polinomios

$$P=\bar{1}x^4+\bar{2}x \quad \text{y} \quad P_1=\bar{1}x^2+\bar{2}x.$$

Para $x=\bar{1}$, $P(\bar{1})=P_1(\bar{1})=\bar{0}$; para $x=\bar{0}$, $P(\bar{0})=P_1(\bar{0})=\bar{0}$, y para $x=\bar{2}$, $P(\bar{2})=P_1(\bar{2})=\bar{2}$.

[1] Si A no tiene raíz, el teorema está, evidentemente, demostrado.

Toman el mismo valor para todo valor de la variable y definen, por tanto, la misma función. Esta circunstancia realza el valor del teorema siguiente.

TEOREMA.—*Si el cuerpo* K *contiene una infinidad de elementos, dos polinomios* P *y* Q *que definan la misma función son iguales (es decir, tienen los mismos coeficientes).*

En efecto, el polinomio $R(x)=P(x)-P_1(x)$, cuyo grado no puede sobrepasar al de P o de P_1, se anula para más de n valores, lo que es imposible en virtud del teorema de la página anterior; por consiguiente, solo puede ser el polinomio nulo.

Este teorema se aplica en particular sobre los cuerpos Q, R y C: la igualdad numérica implica la igualdad formal. Se dice con frecuencia que los dos polinomios son *idénticos.*

Según el resultado precedente, para asegurar la identidad de dos polinomios de grado n, basta escribir que toman los mismos valores para $n+1$ valores diferentes de la variable.

EJERCICIOS:

1. Un polinomio admite como resto k o l cuando se divide por $x-a$ o por $x-b$. Hállese el resto de su división por el producto $(x-a)(x-b)$. Generalícese al caso de tres factores, llamando m el resto de la división por $x-c$.

2. Sobre el cuerpo Z/7, determínese un polinomio $P(x)=ax^2+bx+c$ tal que $P(1)=\bar{4}$, $P(\bar{2})=\bar{0}$, $P(\bar{0})=\bar{1}$.

3. En el anillo K[x], donde K=Z/5, hállese otro polinomio que defina la misma función que el polinomio $\bar{1}x^2-\bar{2}x+\bar{4}$.

4. Si P(x) y Q(x) definen la misma función sobre un cuerpo finito, K, de elementos $a_1, a_2, \ldots, a_n$, pruébese que

$$P(x)-Q(x)=(x-a_1)(x-a_2)\ldots(x-a_n)R(x),$$

donde R(x) es un polinomio cualquiera de K[x].

5. Demuéstrese que el teorema sobre la división por $x-a$ es válido si consideramos el anillo A[x] de los polinomios cuyos coeficientes pertenecen a un dominio de integridad, A. Generalícese el teorema sobre el número máximo de raíces (ceros) de un polinomio.

6. Pruébese que el polinomio $\bar{1}x^2+\overline{14}$, cuyos coeficientes pertenecen al anillo Z/15, tiene cuatro raíces. Explíquese por qué esto no contradice el resultado del ejercicio anterior.

7. Calcúlese el resto de la división de $(\cos a+x \operatorname{sen} a)^n$ por x^2+1.
(*Resp.:* $\cos na+x \operatorname{sen} na$.)

8. Sobre R, descompóngase $P=x^5-209x+56$ en un producto de dos factores, sabiendo que admite dos raíces a y b, inversas una de la otra (P debe ser divisible por $x^2-\lambda x+1$. Efectuando la división se deduce que $\lambda=4$).

16. División según las potencias crecientes.—El algoritmo de la división se ha obtenido hallando en una clase de congruencia una sucesión de polinomios de grados decrecientes. Se puede aplicar una idea análoga tratando de hallar ahora polinomios (ordenados según las potencias crecientes) que comiencen por términos de grados más elevados cada vez. Sea A un polinomio; busquemos en su clase un polinomio $A_1 = A - Bq_0$ que carezca de término constante. Esto es posible si, y solo si, el término constante de B, b_0, es distinto de cero, lo que se supone expresamente. Basta, en efecto, elegir $q_0 = \frac{a_0}{b_0}$, donde a_0 es el coeficiente constante de A, para estar seguros de que el polinomio A_1 comienza por un término de primer grado al menos; sea $a'x$ este término, con a' eventualmente nulo. Se forma a continuación $A_2 = A_1 - Bq_1$ con el fin de suprimir el término de primer grado en la diferencia; para ello es suficiente tomar $q_1 = \frac{a'x}{b_0}$ y A_2 comienza por un término que es al menos de segundo grado; de A_2 se pasa a A_3, etc.

Consideremos de esta forma la sucesión de polinomios A, A_1, A_2, ..., $A_{k+1} = A_k - Bq_k$; forzosamente se presenta una de estas dos circunstancias:

1.ª si A es un múltiplo de B, uno de estos polinomios es nulo y termina la sucesión;

2.ª si A no es múltiplo de B, la operación puede proseguirse indefinidamente; después de $h+1$ operaciones[1] habremos formado un polinomio A_{h+1} congruente con A y que comienza por un término de grado $h+1$ por lo menos. Tal polinomio no es único en una clase de congruencia; pero, sumando las igualdades que definen A_1, A_2, ..., A_{h+1}, se obtiene:

$$A_{h+1} = A - B(q_0 + q_1 + \ldots + q_h)$$

y el polinomio $Q = q_0 + q_1 + \ldots + q_h$ es de grado inferior o igual a h. Se puede, pues, escribir, en todos los casos

$$A = BQ + x^{h+1}R, \qquad \text{con} \qquad \text{grado de } Q \leqslant h.$$

Demostremos que la desigualdad, grado de $Q \leqslant h$, implica la unicidad de Q y de R. Si se tuviera

$$A = BQ + x^{h+1}R = BQ' + x^{h+1}R'$$

[1] Algunas pueden no modificar el polinomio; p. ej., $A_2 = A_1$ si $q_1 = 0$, es decir, $a' = 0$.

con grado de Q o $Q' < h$, se deduciría que $B(Q - Q') = x^{h+1}(R' - R)$. Para $Q \neq Q'$ el primer miembro comenzaría por un término $\alpha x^{h'}$ con $h' \leqslant h$ (ya que $b_0 \neq 0$), lo que es imposible vista la forma del segundo miembro.

TEOREMA.—*Sean* A *y* B *dos polinomios cualesquiera, y el último tal que* $B(0) \neq 0$. *Es posible encontrar un polinomio* Q *y uno solo que satisfaga a*

$$A = BQ + x^{h+1}R, \quad \text{con} \quad \text{grado de } Q \leqslant h,$$

donde h es un entero natural arbitrario.

El polinomio Q se llama cociente de A por B, de orden *h*, dispuesto según las potencias crecientes de *x*. Naturalmente, si A es múltiplo de B, se tiene $R = O$ si *h* se elige lo bastante grande. La operación se efectúa adoptando la disposición clásica.

EJERCICIOS:

1. Efectúese la división en potencias crecientes, de orden *n*, de $1 - abx^2$ por $1 - (a+b)x + abx^2$.
Resp.:

$$Q = 1 + (a+b)x + (a^2+b^2)x^2 + \ldots + (a^n+b^n)x^n;$$
$$R = (a^{n+1}+b^{n+1})x^{n+1} - ab(a^n+b^n)x^{n+2}.$$

2. Divídase 1 por $(1-x)^2$ según las potencias crecientes de *x*. Precísese la expresión general del cociente y de los dividendos parciales. Dedúzcase de lo anterior una expresión sencilla de

$$P = 1 + 2x + 3x^2 + \ldots + nx^{n-1}.$$

IDEALES DE POLINOMIOS. M. C. D. Y M. C. M.

17. Noción de ideal.—El estudio de las clases de congruencia con relación a un polinomio $A \in K[x]$ ha puesto en evidencia la clase particular formada por los múltiplos de A. Constituye un subconjunto $\mathcal{J}(A) \subseteq K[x]$ que posee las tres propiedades siguientes:

1.ª contiene el polinomio nulo;

2.ª si A_1 y A_2 son dos elementos de $\mathcal{J}(A)$, la diferencia $A_1 - A_2$ pertenece también a $\mathcal{J}(A)$;

3.ª si A_1 pertenece a $\mathcal{J}(A)$, cualquiera que sea $P \in K[x]$, el polinomio PA_1 pertenece a $\mathcal{J}(A)$.

Recíprocamente, consideremos un subconjunto $\mathcal{J}$ que verifique las

tres propiedades anteriores. Por definición, tal subconjunto se llama *ideal* de polinomios y goza de la propiedad enunciada en el teorema siguiente.

TEOREMA.—*Todo ideal* $\mathcal{J}$ *de* K[x] *está formado por el conjunto de los múltiplos de un mismo polinomio* A.

En efecto, si $\mathcal{J}$ no contiene más que O se puede decir que está formado por los múltiplos de este polinomio; en caso contrario, por contener un polinomio no nulo, contendrá una infinidad (según 3) y a cada uno de ellos se hace corresponder su grado, que es un número positivo o nulo. El conjunto de estos enteros admite un límite inferior (entero) n_1, con seguridad alcanzado por al menos un polinomio A del conjunto. Demostremos que todo polinomio B de $\mathcal{J}$ es múltiplo de A. En efecto, la división de B por A da

$$B - QA + R, \qquad \text{con} \qquad \text{grado de } R < \text{grado de } A, \qquad \text{o bien } R = O.$$

Puesto que $A \in \mathcal{J}$, $QA \in \mathcal{J}$ (según 3) y $R = B - QA \in \mathcal{J}$ (según 2); pero entonces la hipótesis «grado de R < grado de A» es incompatible con la elección de A. Por consiguiente, debemos suponer R = O, y B es múltiplo de A.

OBSERVACIONES.—1.ª La demostración anterior enseña que, salvo un factor constante, existe un solo polinomio de grado mínimo.

2.ª Cuando $n_1 = 0$, A es una constante arbitraria y el ideal se identifica con K[x]. Hemos obtenido, pues, dos casos de degeneración: A = O, con lo que $\mathcal{J}$ se reduce a O; A = constante, de modo que $\mathcal{J}$ comprende todo el conjunto K[x].

3.ª En cualquier caso, se dice que el polinomio A, cuya existencia acabamos de demostrar, *engendra* el ideal.

4.ª La definición precedente es la particularizacion a los anillos de polinomios de la definición general dada en la página 94.

EJEMPLO.—Sean $A_1, A_2, \ldots, A_n$, n polinomios determinados de K[x]. Demostremos que el conjunto de los polinomios

$$P = X_1A_1 + X_2A_2 + \ldots + X_nA_n,$$

donde $X_1, X_2, \ldots, X_n$ designan polinomios arbitrarios de K[x], constituye un ideal.

En efecto, el conjunto considerado contiene: el polinomio nulo para $X_1 = \ldots = X_n = O$; la diferencia de P y de $P' = X_1'A_1 + \ldots + X_n'A_n$, ya que

$$P - P' = (X_1 - X_1')A_1 + \ldots + (X_n - X_n')A_n,$$

y un múltiplo cualquiera de P, puesto que

$$PQ = (X_1Q)A_1 + \ldots + (X_nQ)A_n.$$

Tenemos, pues, un ideal, que podemos decir engendrado por A_1, A_2, ..., A_n, y al cual pertenecen también estos polinomios. El teorema precedente demuestra que el mismo ideal puede a su vez engendrarse por un polinomio único, el de grado mínimo.

18. M. C. D. Polinomios primos entre sí.—La discusión del carácter propio o degenerado del ideal $\mathcal{J}$ engendrado por n polinomios A_1, A_2, ..., A_n permite introducir ciertas nociones fundamentales relativas a este conjunto de polinomios:

1.° El caso $\mathcal{J} = 0$, exige $A_1 = A_2 = \ldots = A_n = 0$ y puede excluirse sin inconveniente.

2.° En el caso $\mathcal{J} \subset K[x]$ (sentido estricto) existe un polinomio D (grado de $D \geqslant 1$) que engendra a $\mathcal{J}$. En particular, A_1, A_2, ..., A_n son múltiplos de D, o también D es un divisor común de todos estos polinomios. Naturalmente, todo divisor de D es también divisor común. Precisemos la propiedad recíproca.

TEOREMA.—*Todo divisor común a n polinomios divide al polinomio* D *que engendra el mismo ideal que los polinomios considerados.*

Es suficiente observar que todo divisor común a A_1, A_2, ..., A_n divide a todos los polinomios del ideal engendrado por A_1, A_2, ..., A_m, y, por tanto, a D en particular, que es uno de estos polinomios.

Definición.—El polinomio D, divisor común de mayor grado de los n polinomios, se llama *máximo común divisor* (M. C. D.) de estos polinomios.

3.° En el caso $\mathcal{J} = K[x]$, D es una constante arbitraria. Las propiedades precedentes continúan siendo válidas, aunque adquieren un significado muy diferente. El hecho de que D sea un divisor común apenas dice nada, pero la recíproca demuestra que los polinomios A_1, A_2, ..., A_n no tienen divisor común en sentido estricto. Se llaman *primos entre sí* (en su conjunto).

Como resumen enunciamos la siguiente proposición:

TEOREMA.—*Dados n polinomios no todos nulos, o admiten un M. C. D., o son primos entre sí, según que el ideal que engendran sea propio o degenerado.*

Teorema de Bezout.—*Una condición necesaria y suficiente para que n polinomios sean primos entre sí en su conjunto es que existan*

n polinomios $X_1, X_2, \ldots, X_n$ *tales que*

$$X_1A_1 + X_2A_2 + \ldots + X_nA_n = 1 \qquad \text{(elemento unidad de K)}.$$

La existencia de los polinomios $X_1, X_2, \ldots, X_n$ implica que el ideal es degenerado; inversamente, si este se confunde con $K[x]$, el número 1 es combinación de $A_1, A_2, \ldots, A_n$, luego los polinomios $X_1, X_2, \ldots, X_n$ existen.

Propiedades del M. C. D.—1.ª *Si multiplicamos* $A_1, A_2, \ldots, A_n$ *por* B, *el M. C. D. queda multiplicado por* B, *y, recíprocamente, si dividimos* $A_1, A_2, \ldots, A_n$ *por uno de sus divisores comunes, el M. C. D. queda dividido por este divisor.*

2.ª *Condición necesaria y suficiente para que* $A_1, A_2, \ldots, A_n$ *admitan un divisor común,* D, *como M. C. D. es que los cocientes de* $A_1, A_2, \ldots, A_n$ *por* D *sean primos entre sí (en su conjunto).*

OBSERVACIÓN.—Hemos supuesto dado el cuerpo K y se ha razonado sobre el anillo $K[x]$. Sin embargo, si elegimos n polinomios, pueden pertenecer a anillos correspondientes a cuerpos cada vez menos vastos, $K_1 \subset K_2 \subset K_3 \ldots$ El menor anillo $K_1[x]$ corresponde al mínimo cuerpo que contiene a todos los coeficientes de los polinomios dados [1]. Para n polinomios, el carácter de ser primos entre sí es independiente del cuerpo, $K_1, K_2, K_3, \ldots$, elegido. En efecto, la igualdad de Bezout

$$X_1A_1 + X_2A_2 + \ldots + X_nA_n - 1$$

es verdadera o falsa sobre uno cualquiera de los anillos considerados: Si es cierta sobre $K_2[x]$, p. ej., lo será sobre uno cualquiera de los siguientes, según la inclusión $K_2 \subset K_3$. También lo es sobre $K_1[x]$, pues, en otro caso, los polinomios no primos sobre K_1 no podrían serlo sobre K_2. Finalmente, si la igualdad puede verificarse, lo será sobre $K_1[x]$. Resulta de ello también que el M. C. D. de n polinomios pertenece a $K_1[x]$ y no varía cualquiera que sea la ampliación que se haga del cuerpo base.

19. Caso de dos polinomios.—El caso $n = 2$ es particularmente interesante. Dos polinomios A y B cuyo M. C. D. sea igual a 1 se llaman primos entre sí, pudiendo suprimirse la expresión «en su conjunto». En

[1] Si los coeficientes son reales, este será con frecuencia un subcuerpo de R. Así, p. ej., para $A = (1+\sqrt{2})x^2 + 2x - \frac{1}{\sqrt{2}}$ y $B = x^3 - 3\sqrt{2}x + 5$, este será el cuerpo $a + b\sqrt{2}$, con a y b racionales.

el caso $n \geqslant 3$, indica que queda excluida la consideración de los polinomios dos a dos. Precisemos el estudio del caso $n = 2$.

TEOREMA.—*Si un polinomio* A *es primo con cada uno de los polinomios* B *y* C, *es primo con su producto* BC.

Según el teorema de Bezout, tenemos los polinomios X_1, X_2, X'_1, X'_2 tales que

$$X_1A + X_2B = 1 \quad \text{y} \quad X'_1A + X'_2C = 1.$$

Multiplicando miembro a miembro, la igualdad obtenida se escribe en la forma

$$X''_1A + X''_2BC = 1 \qquad (X''_2 = X_2X'_2).$$

Por el teorema de Bezout se ve que A y BC son primos entre sí.

CONSECUENCIAS.—*Si* A *y* B *son primos entre sí,* A^n *y* B^p *también lo son cualesquiera que sean los enteros* n *y* p.

Teorema de Gauss.—*Todo polinomio* A *que divide a un producto de dos factores* BC *y es primo con uno de ellos,* B, *divide al otro,* C.

El teorema de Bezout se escribe

$$X_1A + X_2B = 1,$$

de donde, multiplicando por C,

$$(CX_1)A + X_2BC = C.$$

El polinomio A divide a BC y, por consiguiente, divide al primer miembro, es decir, a C.

Disposición práctica.—Se reduce a determinar el polinomio D de menor grado que engendra el ideal $\mathcal{J}: X_1A + X_2B$. Partiendo de A y de B se busca en $\mathcal{J}$ una sucesión de polinomios de grados cada vez menores, y forzosamente se debe llegar a D después de un número finito de operaciones. Si grado de $B \geqslant$ grado de A, tendremos $B = QA + A_1$, con grado de $A_1 <$ grado de A y $A_1 \in \mathcal{J}$; A y A_1 engendran $\mathcal{J}$ y son de grados respectivamente inferiores a B y a A; continuando así, $A = Q_1A_1 + A_2$, con grado de $A_2 <$ grado de A_1, etc. Si A y B no son primos entre sí, necesariamente se llega a una división con resto nulo: el resto anterior engendra el ideal y representa el M. C. D. Si se llega a un resto de grado nulo, $\mathcal{J}$ es degenerado, y A y B, primos entre sí.

El lector habrá reconocido en este proceso el algoritmo de Euclides, estudiado en aritmética en el caso de los números enteros.

EJEMPLO.—Siendo K el cuerpo de los racionales, calcúlese el M. C. D. de los polinomios $3x^4+9x^3-3x^2-12x-9$ y $3x^3+10x^2+2x-3$.

La primera división da como cociente $x+1$ y como resto $-x^3-5x^2-9x-9$. Para simplificar la escritura, evitando las fracciones, se puede multiplicar este resto por -3. Se obtiene así:

$$3x^3+15x^2+27x+27;$$

la segunda división da entonces 1 como cociente y $5x^2+25x+30$ como resto. Para mayor sencillez se considera el polinomio x^2+5x+6. La tercera división da $3x-5$ como cociente y $9x+27$ como resto; tomamos $x+3$. La última división resulta exacta y el M. C. D. normalizado es $x+3$.

EJERCICIOS:

1. Aplíquese el algoritmo de Euclides a los siguientes polinomios:

a) $A=x^5-x^4+2x^3+1$, $B=x^5+x^4+2x^2-1$;
b) $A=x^4+2x^3-11x^2-12x+36$, $B=4x^3+6x^2-22x-12$;
c) $A=x^6-2x^5+x^4-x^2+2x-1$, $B=x^5-3x^3+x^2+2x-1$.

2. Calcúlese el M.C.D. de $x^{99}+1$ y $x^{45}+1$.

3. Demuéstrese que si A y B son primos entre sí, A+B y AB también lo son.

4. ¿Qué relación deben satisfacer a y b para que los dos polinomios x^7-a y x^5-b admitan un M.C.D. de primer grado? Determínese.

5. Si se conocen dos polinomios S_0 y T_0 que verifican la igualdad de Bezout con relación a dos polinomios primos entre sí, A y B, demuéstrese cómo se pueden hallar a partir de ellos todos los demás polinomios S y T que satisfacen dicha igualdad.

Concluyase de lo anterior que existe un par y uno solo de polinomios S y T tales que

$$\text{grado de } S < \text{grado de } B \quad \text{y} \quad \text{grado de } T < \text{grado de } A.$$

6. Pruébese que, en el algoritmo de Euclides, todo polinomio de la sucesión $A_1, A_2, \ldots, A_n$ puede ponerse en la forma

$$A_i = S_iA + T_iB.$$

Razónese por recurrencia y dedúzcase de ello el teorema de Bezout.

7. En el anillo Q[x] determínense los polinomios S y T del teorema de Bezout suponiendo $A=x^m$ y $B=(1-x)^p$ y, además, grado de $S<p$, grado de $T<m$.

8. El M.C.D. de $x^m - a^m$ y de $x^p - a^p$ es $x^h - a^h$, donde h es el M.C.D. de m y p.

20. Mínimo común múltiplo.—Razonemos sobre n polinomios $A_1, A_2, \ldots, A_n$. Siempre admiten múltiplos comunes; p. ej., los polinomios de la forma $M = PA_1A_2 \ldots A_n$, donde P es arbitrario. Estudiemos con más detalle el conjunto de estos múltiplos.

TEOREMA.—*El conjunto de los múltiplos comunes a n polinomios constituye un ideal del anillo de los polinomios.*

En efecto, O es evidentemente múltiplo común; si M_1 y M_2 lo son, igual sucede con $M_1 - M_2$ y PM_1, como se demuestra fácilmente. El ideal obtenido, $\mathcal{J}$, no queda ciertamente reducido a O si los polinomios $A_1, A_2, \ldots, A_n$ son todos distintos de O, porque $A_1A_2 \ldots A_n \in \mathcal{J}$. Por el contrario, si al menos un A_i es nulo, lo serán todos sus múltiplos e $\mathcal{J} = O$. Descartemos este caso. Si el ideal $\mathcal{J}$ se confunde con K[x], las igualdades

$$P = P_1A_1 = P_2A_2 = \ldots = P_nA_n$$

son posibles cualquiera que sea P; por consiguiente, todos los A_i son constantes. En consecuencia, si los polinomios dados son diferentes de O y de grado superior a cero, el ideal no es degenerado. Existe un polinomio, M, de mínimo grado, cuyos múltiplos constituyen el conjunto de múltiplos comunes a $A_1, A_2, \ldots, A_n$. Este polinomio es el M. C. M. de los polinomios dados.

CASO $n = 2$.—En este caso se puede precisar la determinación del M. C. M.

TEOREMA.—*El M. C. M. de dos polinomios, A y B, es igual al cociente de su producto por su M. C. D.*

Si M es el M. C. M. buscado, se tiene que $AB = KM$ con $M = P_1A = Q_1B$. Sustituyendo, se comprueba que K es un divisor común de A y de B. Se obtendrá el M. C. M. tomando para K el M. C. D. de A y B, con lo que el teorema queda demostrado.

EJERCICIOS:

1. Hállense el M.C.D. y el M.C.M. de los polinomios

$$A = x^8 + x^7 - 2x^6 - 3x^5 + 3x^3 + 2x^2 - x - 1,$$
$$B = 8x^7 + 7x^6 - 12x^5 - 15x^4 + 9x^2 + 4x - 1.$$

2. Demuéstrese que el M.C.M. de tres polinomios A, B, C es también el M.C.M. de C y del M.C.M., D, de A y B.

3. Hállese el M.C.M. de los tres polinomios x^3-7x-6, x^3-3x+2, x^3-4x^2+3x.

4. Indíquese la forma general de los polinomios A y B que admiten como M.C.D. su diferencia $D=A-B$.

Determínense estos polinomios, conociendo su M.C.M., $4x^3-10x^2-6x$ (varias soluciones).

NOCIONES SOBRE ELIMINACION

21. Polinomios no primos.—Sean A y B dos polinomios de K[x]. Tratamos de hallar una condición necesaria y suficiente para que admitan divisores comunes en sentido estricto; es decir, un M. C. D. de por lo menos primer grado. Si es efectivamente de primer grado, en virtud del teorema de la página 227, se obtiene una condición necesaria y suficiente para que los dos polinomios admitan una raíz común. La condición buscada es, por definición, el resultado de *eliminar* x entre los polinomios $A(x)$ y $B(x)$.

Los casos elementales que se estudian a continuac ón justifican el término «eliminar».

TEOREMA.—*Una condición necesaria y suficiente pa a que los polinomios* A *y* B *de* K[x], *supuestos efectivamente de grados* n *y* p, *no sean primos entre sí, es que los polinomios* A, xA, x^2A, ..., x^{p-1}A *y* B, xB, x^2B, ..., x^{n-1}B *sean linealmente dependientes sobre* K.

En efecto, es condición necesaria y suficiente para que A y B no sean primos que su M. C. M. sea de grado inferior al del producto AB. Se tiene:

$$M-P_1A=Q_1B$$

y los grados de los polinomios P_1 y Q_1 son respectivamente inferiores a p y a n, y recíprocamente. La condición buscada equivale, pues, a que se verifique la igualdad

$$P_1A+(-Q_1)B=O \quad \text{con} \quad \text{grado } P_1<p, \quad \text{grado de } Q_1<n.$$

Poniendo

$$P_1=\lambda_0+\lambda_1x+\ldots+\lambda_{p-1}x^{p-1}, \qquad Q_1=\mu_0+\ldots+\mu_{n-1}x^{n-1},$$

aquella se escribe

$$\lambda_0A+\lambda_1(xA)+\ldots+\lambda_{p-1}(x^{p-1}A)+\mu_0B+\mu_1(xB)+\ldots+\mu_{n-1}(x^{n-1}B)=0$$

que es equivalente a la enunciada condición de dependencia.

CONSECUENCIAS.—Los polinomios A, xA, ..., $x^{p-1}A$, B, xB, ..., $x^{n-1}B$ constituyen $n+p$ polinomios de grados inferiores o iguales a $n+p-1$, luego pertenecen a un subespacio de $n+p$ dimensiones del espacio $K[x]$. Puede suceder, por consiguiente, que estos polinomios sean independientes, lo que corresponde a A y B primos entre sí. Si se quiere que sean dependientes, basta asegurar la existencia de una relación y una sola entre los coeficientes de los polinomios dados [1]. Antes de examinar casos particulares importantes precisemos la condición precedente.

TEOREMA.—*Si el rango de los $n+p$ polinomios definidos anteriormente es igual a $n+p-1$, el M. C. D. de A y B es de primer grado.*

Este M. C. D. existe, y si es k su grado, se podrán hallar dos polinomios, P_1 y Q_1, de grados respectivos $p-k$ y $n-k$, tales que

$$AP_1+BQ_1=0;$$

es decir, que A, xA, ..., $x^{p-k}A$ y B, xB, ..., $x^{n-k}B$ son dependientes. En la relación obtenida, el coeficiente de $x^{p-k}A$, igual al último coeficiente de P_1, es con seguridad distinto de cero y, por tanto, $x^{p-k}A$ es combinación lineal de los otros polinomios. Si k es superior a 1, esto contradice la hipótesis hecha acerca del rango. En efecto, para $k=2$, p. ej., se encontraría que $x^{p-2}A$ y (después de multiplicar por x) $x^{p-1}A$ son combinaciones lineales de los $n+p-2$ polinomios restantes. Debemos, pues, suponer $k=1$, lo que demuestra el teorema.

Resultante de dos polinomios de segundo grado.—Apliquemos el resultado general al caso de dos polinomios de segundo grado:

$$A=ax^2+bx+c, \qquad B=a'x^2+b'x+c' \qquad (a \text{ y } a' \neq 0).$$

Ello nos lleva a escribir que cuatro vectores de un espacio de cuatro dimensiones son dependientes, lo que se hace mediante los métodos de la página 139.

Adoptando por base x^3, x^2, x, x^0, se escribe la matriz de los componentes de xA, A, xB, B:

$$\begin{bmatrix} a & b & c & 0 \\ 0 & a & b & c \\ a' & b' & c' & 0 \\ 0 & a' & b' & c' \end{bmatrix}.$$

[1] En general, se obtiene una condición y solo una escribiendo que un vector de un espacio de k dimensiones pertenece a un subespacio de $k-1$ dimensiones (véase pág. 353).

Transformándola, se obtiene:

$$\begin{bmatrix} a & b & c & 0 \\ 0 & a & b & c \\ 0 & ab'-ba' & ac'-ca' & 0 \\ 0 & a'c-c'a & b'c-c'b & 0 \end{bmatrix}$$

y escribimos que las dos últimas líneas son proporcionales:

$$(ab'-ba')(bc'-cb')-(ac'-ca')^2=0.$$

La transformación supone que c (o c') es diferente de cero. El lector se asegurará de que el resultado continúa siendo válido si uno de estos coeficientes es nulo.

El primer miembro de la relación precedente es la *resultante* de los polinomios A y B.

OTRO EJEMPLO.—Si $A=a_0+\ldots+a_nx^n$ es cualquiera y $B=ax+b$ de primer grado, la condición buscada se obtiene escribiendo que la raíz $x=-b/a$ de B es raíz de A; o sea, en forma entera,

$$a_0a^n-a_1a^{n-1}b+\ldots+(-1)^na_nb^n=0.$$

Método de reducción.—En la práctica, el empleo de la matriz es bastante incómodo si los coeficientes no son numéricos. Se prefiere reemplazar los polinomios A y B por otros de grado inferior y que admitan, en general, el mismo M. C. D. que A y B. Los casos de excepción pueden estudiarse aparte. Repitiendo cierto número de veces el método se llega a dos polinomios de segundo grado y se escribe su resultante. De todas formas, para polinomios de grados elevados, el proceso resulta muy laborioso.

EJEMPLO. Elimínese x entre los polinomios

$$A=ax^3+bx^2+cx+d, \qquad B=a'x^3+b'x^2+c'x+d' \qquad (a,\ a'\neq 0).$$

Los polinomios

$$a'A-aB=(a'b-ab')x^2+(a'c-ac')x+(a'd-ad')$$

y

$$\frac{d'A-dB}{x}=(d'a-da')x^2+(d'b-db')x+(d'c-dc')$$

son de segundo grado. Con certeza admiten como divisor común el

M. C. D. de A y B, salvo si este contiene el factor x, lo que exigiría $d=d'=0$. La cuestión queda así reducida a escribir la resultante de las dos ecuaciones que preceden. Por otra parte, es nula para $d=d'=0$ y nos da una condición necesaria y suficiente.

EJERCICIOS:

1. Elimínese x entre los polinomios

$$P=x^3+px+q, \qquad Q=x^2+ax+b.$$

Resp.: $(q+ab)^2+(pb-b^2-aq)(a^2-b+p)=0$.

2. Determínese m para que los dos polinomios

$$P=x^3+mx-6, \qquad Q=x^2+mx-2$$

tengan una raíz común. *Resp.:* $(m+1)(m^2+7)=0$.

3. ¿Para qué valores de λ tienen una raíz común los siguientes polinomios:

$$A=x^3-\lambda x+2 \qquad y \qquad B=x^2+\lambda x+2?$$

4. La misma cuestión para los polinomios

$$A=x^3-2\lambda x+\lambda^3 \qquad y \qquad B=x^2+\lambda^2-2.$$

Resp.:

$$\lambda=1, \quad \lambda=\tfrac{1}{2}\left(-2+\sqrt{2}\pm\sqrt{4\sqrt{2}-2}\right)$$

y

$$\lambda=\tfrac{1}{2}\left(-2-\sqrt{2}\pm i\sqrt{4\sqrt{2}+2}\right)$$

5. Demuéstrese el recíproco del teorema de la página 238: si el M.C.D. de A y B es de primer grado, el rango de los $n+p$ polinomios considerados es igual a $n+p-1$.

6. Generalícese la propiedad precedente demostrando que es condición necesaria y suficiente, para que dos polinomios A y B tengan un M.C.D. de grado h, que el rango de los $n+p$ polinomios definidos en la página 237 sea igual a $n+p-h$.

EJERCICIOS SOBRE EL CAPITULO XI

11-1. *a)* Efectúese la división de $P(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n$ por $x-\alpha$. Compruébese que los coeficientes del cociente $Q(x)=b_0x^{n-1}+b_1x^{n-2}+\ldots+b_{n-1}$ son tales que

$$b_0=a_0, \qquad b_i=\alpha b_{i-1}+a_i,$$

siendo el resto el elemento b_n de la sucesión anterior.

b) Puesto que $b_n=f(\alpha)$, de ello deducimos un método de cálculo numé

rico de $f(\alpha)$ (método de Horner). En la práctica se emplea el esquema que sigue:

$$\begin{array}{c|cccc} \alpha & a_0 & a_1 & a_2 & \cdots & a_n \\ \hline & b_0 & b_1 & \cdots & \end{array}$$

donde cada término de la segunda línea se obtiene sumando al término correspondiente de la primera el producto por α del término precedente de la segunda. Aplíquese este método en el ejercicio que sigue.

11-2. *Raíces racionales de una ecuación.*

1.º El cuerpo base es Q. Mediante operaciones evidentes se consigue que la ecuación

$$a_0x^n + a_1x^{n-1} + \ldots + a_n = 0 \qquad \text{(E)}$$

tenga por coeficientes números enteros relativos primos entre sí en su conjunto.

Demuéstrese el siguiente teorema: si la ecuación (E) admite una raíz igual a la fracción irreducible p/q, p divide necesariamente a a_n y q divide a a_0.

2.º Hállense las raíces racionales de la ecuación

$$4x^4 - 10x^3 - 38x^2 + 70x + 43 = 0;$$

para ensayar las posibles raíces se utilizará el esquema de Horner.

3.º Lo mismo para

$$10x^5 - 81x^4 + 90x^3 - 102x^2 + 80x - 21 = 0.$$

CAPITULO XII

DESCOMPOSICION EN FACTORES PRIMOS FRACCIONES RACIONALES

FACTORES PRIMOS. TEOREMA FUNDAMENTAL

1. Polinomio irreducible sobre un cuerpo.—Las teorías desarrolladas en el capítulo anterior han demostrado la analogía de las propiedades del anillo K[x] y del anillo de los enteros relativos. Vamos a extender a K[x] la noción de factor primo y la descomposición clásica de la aritmética.

Definición.—Un polinomio P del anillo K[x] se dice *irreducible* o *primo* sobre K cuando no puede descomponerse en producto de polinomios de grado inferior que pertenezcan a K[x]. En caso contrario, el polinomio dado se llama reducible o descomponible sobre K, y se puede escribir:

$$P = P_1 \cdot P_2 \quad \text{con} \quad \text{grado de } P_1 \text{ o } P_2 < \text{grado de } P.$$

La condición acerca de los grados se interpreta en sentido estricto; por tanto, P_1 y P_2 no pueden ser constantes. Se sabe que las constantes, que constituyen los elementos unitarios del anillo, dividen a todos los polinomios. En una descomposición de la forma $P = P_1 \cdot P_2$ se puede adoptar para P, P_1, P_2 polinomios normalizados en el sentido indicado en la página 219, o bien razonar con las clases de congruencia respecto de la multiplicación por una constante y no considerar como distintas dos descomposiciones cuyos factores sean proporcionales.

EJEMPLOS.—1.° Sobre un cuerpo cualquiera los binomios de primer grado, $x - a$, son polinomios primos.

2.° En R, un trinomio de segundo grado $x^2 + px + q$ tal que $p^2 - 4q < 0$ es un polinomio primo, ya que toda descomposición comprendería necesariamente dos factores de primer grado y el trinomio debería admitir raíces reales, en contra de la hipótesis. En cambio, sobre C, un trinomio es descomponible sin excepción en la forma $(x - \alpha)(x - \beta)$, ya que una ecuación de segundo grado admite siempre dos raíces.

OBSERVACIÓN IMPORTANTE.—El carácter reducible o irreducible de un polinomio P *dado* depende esencialmente del cuerpo base elegido. Los coeficientes de P son, evidentemente, elementos de un cuerpo K, pero muchas veces pertenecen a subcuerpos más o menos amplios:

$$K_1 \subset K_2 \subset \ldots \subset K.$$

La descomposición de P depende en definitiva de que se elija K_1, K_2 o K como cuerpo base. Indiquemos algunos ejemplos.

1.° $P = x^2 + 3$ es indescomponible en el cuerpo de los racionales, en el de los números reales, e incluso en todo cuerpo ordenado, ya que en estos un cuadrado es siempre positivo. Pero, en C, tenemos $x^2 + 3 = (x + i\sqrt{3})(x - i\sqrt{3})$.

2.° $P = x^2 - 2$ es descomponible en el cuerpo de los números reales:

$$(x^2 - 2) = (x + \sqrt{2})(x - \sqrt{2}).$$

No lo es en el de los racionales: la descomposición comprendería dos factores de primer grado, y la ecuación $x^2 - 2 = 0$ tendría dos raíces racionales, lo que es imposible dado que no existe ninguna fracción cuyo cuadrado sea igual a 2.

3.° $P = x^3 - 2$ es primo en Q; en cambio, se tiene:

en R: $$P = (x - \sqrt[3]{2})(x^2 + x\sqrt[3]{2} + \sqrt{4})$$

y en C: $$P = (x - \sqrt[3]{2})\left(x + \frac{1 - i\sqrt{3}}{\sqrt[3]{4}}\right)\left(x + \frac{1 + i\sqrt{3}}{\sqrt[3]{4}}\right).$$

TEOREMA.—*Si un polinomio primo* P *divide a un producto de factores, divide a uno de ellos.*

Para ello basta observar que un polinomio primo es necesariamente primo con todo polinomio al que no divida. En consecuencia, si P|ABC y no divide a A, será primo con A y dividirá a BC (teorema de Gauss); si no divide a B deberá dividir a C. Por consiguientte, divide necesariamente a uno de los tres polinomios A, B o C.

2. Teorema fundamental.—*En el anillo* K[x] *todo polinomio* P *puede descomponerse en un producto que comprenda un elemento unitario y factores irreducibles normalizados.*

La descomposición es única, salvo el orden de los factores.

Si P es una constante o un polinomio irreducible el teorema es evidente. Si P no es irreducible, se tiene $P = P_1P_2$ y los grados de P_1 y P_2 son inferiores al de P. Reiterando el proceso un número finito de veces (como máximo $n-1$ si P es de grado n) se llega necesariamente a factores irreducibles que basta normalizar.

Unicidad.—Sean $P = \lambda p_1 p_2 \dots p_m = \mu q_1 q_2 \dots q_n$ dos descomposiciones. Es necesario primero que $\lambda = \mu$ por la unicidad del término director de P; además, p_1, que divide al producto $q_1 q_2 \dots q_n$, debe dividir a uno de los factores q, y, como se trata de polinomios primos normalizados, $p_1 = q_i$. Se suprime este factor común y se repite el razonamiento. Se ve que cada una de las descomposiciones ha de contener todos los factores de la otra, lo que exige la unicidad.

OBSERVACIÓN.—Los factores $p_1, p_2, \dots, p_m$ no son, necesariamente, distintos. Es cómodo agrupar los factores que sean iguales, lo que da

$$P = \lambda (p_1)^{\alpha} (p_2)^{\beta} \dots (p_j)^{\varepsilon}.$$

CONSECUENCIAS.—Este teorema, que es análogo a otro bien conocido de la aritmética, implica consecuencias idénticas:

1.ª Una vez conocida la descomposición de un polinomio en factores primos es fácil formar la lista de todos sus divisores (definidos salvo un factor constante).

2.ª Para que el polinomio B divida al A es necesario y suficiente que todos los factores primos de B figuren en A con un exponente al menos igual.

3.ª El M. C. D. de dos polinomios se obtiene tomando los factores primos comunes a las dos descomposiciones, afectados cada uno de su menor exponente.

4.ª El M. C. M. de dos polinomios se obtiene tomando todos los factores primos diferentes en ambas descomposiciones, afectado cada uno de su mayor exponente.

TEOREMA.—*Sean dos polinomios* A *y* B *del anillo* K[x]. *Si* A(x) *es primo en* K *y si* A *y* B *tienen una raíz común que pertenece a un supercuerpo* K_1 *de* K, *necesariamente* A *divide a* B.

En efecto, A y B no son primos entre sí en $K_1[x]$. Según la observación hecha en la página 233, tampoco pueden serlo en K[x] y, ya que A es primo en K[x], dividirá necesariamente a B.

Así, p. ej., el polinomio $B = x^5 + 5x^3 - 2x^2 - 10$ de Q[x] es nulo para $x = \sqrt[4]{2}$; el polinomio $A = x^3 - 2$ de Q[x] también se anula; siendo A primo, divide a B, como se comprueba fácilmente.

EJERCICIOS:

1. Demuéstrese que el polinomio $x^3-3n^2x+n^3$ es irreducible en $Q[x]$ si n es un entero relativo.

2. ¿Son reducibles los polinomios $2x^2-1$, $1x^2+1$, y $2x^2+4x+\bar{4}$ en el cuerpo $Z/5$?

3. En $Z/7$, descompóngase $1x^3+2x^2+2x-2$ en factores primos.

3. Descomposición completa.—Se dice que un polinomio P está completamente descompuesto en un cuerpo cuando los factores primos de su descomposición son todos de primer grado. Si P es de grado n, se deben obtener n factores, iguales o distintos; por tanto, tendremos:

$$P=a_n(x-x_1)(x-x_2)\dots(x-x_n).$$

Operando, se debe obtener de nuevo el polinomio dado, lo que conduce a las siguientes *relaciones entre los coeficientes y las raíces:*

$$x_1+x_2+\dots+x_n=-\frac{a_{n-1}}{a_n},$$

$$x_1x_2+x_1x_3+\dots+x_ix_j+\dots+x_{n-1}x_n=\frac{a_{n-2}}{a_n},$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$x_1x_2\dots x_n=(-1)^n\frac{a_0}{a_n}$$

La primera da la suma de las raíces; la segunda, la suma de sus productos binarios; ..., y la última, el producto de todas ellas. Los primeros miembros se llaman polinomios simétricos fundamentales respecto a las n variables.

Admitiendo que todo polinomio simétrico en $x_1, x_2, \dots, x_n$ se expresa racionalmente con ayuda de los polinomios fundamentales (capítulo 13), se deduce como conclusión que tal polinomio depende también, racionalmente, de los coeficientes de la ecuación cuyas raíces son $x_1, x_2, \dots, x_n$.

OBSERVACIÓN.—El cuerpo de los complejos es de descomposición completa, es decir, que los únicos polinomios primos son los binomios de primer grado (véase pág. 251). Las relaciones anteriores son siempre válidas en él. Los siguientes ejercicios deben resolverse exclusivamente en $C[x]$.

EJERCICIOS:

1. Sea $P=2x^3-x^2-7x+\lambda$. Determínese λ para que la suma de dos raíces sea igual a 1.

2. Sea $P=x^3-7x+\lambda$. Determínese λ de manera que una de las raíces sea el doble de la otra.

3. Hállese λ de manera que las raíces del polinomio

$$P=x^3+2x^2-7x+\lambda$$

verifiquen la relación $x_1^2=x_2^2+x_3^2$.

4. Determínense λ y μ para que las raíces del polinomio

$$P=x^4-4x^3-36x^2+\lambda x+\mu$$

estén en progresión aritmética.

EXTENSION ALGEBRAICA DE UN CUERPO

4. Anillo de las clases residuales.—En el capítulo anterior hemos definido las clases de congruencia con relación a un polinomio $M(x)$. Se escribe

$$A \equiv B \text{ (mód. M)}$$

si, y solo si, $A-B$ es divisible por M. El conjunto $K[x]$ queda así dividido en clases disjuntas y, en cada clase, tenemos un polinomio y solo uno de grado inferior al de M, salvo en la clase de los múltiplos de M que contiene a O.

El método que ha permitido definir el anillo Z/n de las clases residuales, a partir del anillo Z, se aplica igualmente al estudio de la congruencia anterior. La adición y la multiplicación en $K[x]$ son estables respecto de la congruencia módulo M y definen una adición y una multiplicación en el conjunto de las clases, operaciones que satisfacen las propiedades asociativa, conmutativa y distributiva.

TEOREMA.—*Las clases residuales definidas módulo* $M(x)$ *en el anillo* $K[x]$ *forman ellas mismas un anillo.*

Este anillo, llamado anillo cociente, se simboliza por $K[x]/M(x)$. Un elemento del mismo es la clase de los polinomios de $K[x]$ que admiten como resto el mismo polinomio $R(x)$ de grado inferior a M. Tal clase se denota por $\bar{A}(x)$, siendo $A(x)$ uno cualquiera de sus representantes; muchas veces se toma $A(x)=R(x)$. El conjunto de los múltiplos de M, comprendido O, forman el elemento nulo $\bar{O}$. Para a elementos

del cuerpo K, el polinomio $A=a$ admite como resto el mismo a. El anillo $K[x]/M(x)$ contiene, pues, el conjunto $\overline{K}$ formado por las clases del tipo $\bar{a}$ y que constituye un subanillo isomorfo a K, del que solo difiere en detalles de notación ($\bar{a}$ en lugar de a) desde un punto de vista formalista. Los resultados anteriores, en los casos de los enteros relativos, números racionales, etc., sugieren la identificación, por convenio, de K con su isomorfo $\overline{K}$. En tales condiciones, K aparece como un subanillo de $K[x]/M(x)$ y este último como un superanillo de K; en particular, en los casos aquí considerados, donde $K[x]/M(x)$ es un cuerpo, se define un supercuerpo de K.

Estas condiciones particularizan para los anillos de polinomios el estudio general expuesto en la página 94. El anillo $K[x]$ es homomorfo a $K[x]/M(x)$.

5. Extensión algebraica de un cuerpo. TEOREMA.—*El anillo* $K[x]/M(x)$ *es un cuerpo si, y solo si,* M *es un polinomio indescomponible.*

Sea A un polinomio no múltiplo de M. Si M no se puede descomponer, será primo con A. Según el teorema de Bezout, es posible hallar una relación de la forma $e=A\cdot B+M\cdot C$ (e, unidad de K), de donde pasando a las clases $\overline{e}=\overline{A}\cdot\overline{B}$, lo que demuestra que toda clase $\overline{A}\neq 0$ admite un inverso; por tanto, el anillo de las clases residuales es un cuerpo.

Si M es descomponible en $M_1\cdot M_2$ de manera no trivial, se tiene $\overline{M}_1\cdot\overline{M}_2=\overline{0}$ y el anillo de las clases residuales admite divisores de cero; por consiguiente, no es un cuerpo.

Definición.—Cuando $M(x)$ no se puede descomponer, el cuerpo de las clases residuales, considerado como supercuerpo de K, es *una extensión algebraica del cuerpo* K.

6. Ejemplo 1.º: el cuerpo de los complejos.—La teoría anterior conduce a una construcción del cuerpo de los números complejos independiente por completo del método expuesto en la parte II. Supongamos construido R y consideremos el anillo $R[x]$; en él, el polinomio x^2+1. es primo y, en consecuencia, el anillo $R[x]/(x^2+1)$ de las clases residuales es un cuerpo. Cada clase residual se representa por el resto común a todos los polinomios de esta clase; es decir, por un polinomio de primer grado $ax+b$, con a y b reales. Este resto se obtiene a partir de un polinomio cualquiera de la clase considerada reemplazando en

él x^2 por -1 cuantas veces sea posible. En efecto, si en el monomio ax^p $(p \geqslant 2)$ se sustituye x^2 por -1 se obtiene $-ax^{p-2}$ tal que

$$ax^p = ax^{p-2}(x^2+1) - ax^{p-2} \qquad \text{o} \qquad ax^p \equiv -ax^{p-2} \text{ mód. } (x^2+1).$$

Efectuando esta sustitución tantas veces como sea posible en cada término del polinomio, no se sale de la clase considerada y se obtiene finalmente el resto $ax+b$. Precisemos las reglas operativas a las que obedecen los elementos así formados que, no se olvide, representan clases.

1.° Suma.— Según el teorema número III (pág. 226), la clase suma de las que tienen por restos $ax+b$ y $a'x+b'$ admite por resto

$$(a+a')x+b+b'.$$

Sobre el conjunto de las clases tenemos, pues, la adición

$$(ax+b)+(a'x+b') = (a+a')x+b+b'$$

la cual conduce a una ley de grupo que no es necesario precisar.

2.° Producto.— Según el teorema número IV (pág. 226), la clase producto de las clases $ax+b$ y $a'x+b'$ contiene el polinomio

$$(ax+b)(a'x+b') = aa'x^2+(ab'+ba')x+bb'.$$

Su resto se obtiene reemplazando x^2 por -1 en la fórmula anterior, y resulta:

$$(ab'+ba')x+bb'-aa',$$

que sobre el conjunto de las clases se escribe:

$$(ax+b)(a'x+b') = (ab'+ba')x+bb'-aa'.$$

En este cálculo x se comporta como un nuevo elemento del cuerpo que obedece a las reglas habituales y cuyo cuadrado sería reemplazable por -1. En consecuencia, el cuerpo $\mathbf{R}[x]/(x^2+1)$ es isomorfo al cuerpo que se obtendría haciendo en R la adjunción de un nuevo elemento i cuyo cuadrado fuera, por definición, *igual* a -1. La existencia y la no contradicción de este cuerpo $\mathbf{R}(i)$ nos viene garantizada por la construcción de su isomorfo $\mathbf{R}[x]/(x^2+1)$, por lo que no emplearemos más el condicional.

TEOREMA.—*Existe una extensión algebraica del cuerpo de los reales, obtenida por la adjunción a* R *de un elemento* i *tal que* $i^2=-1$.

El cuerpo obtenido, R(i), es el de los complejos. En R(i) se pueden descomponer ciertos polinomios no descomponibles en R: el más sencillo es

$$x^2+1=(x+i)(x-i).$$

OBSERVACIÓN.—A partir del anillo Q[x] se construye de manera análoga el cuerpo Q[x]/(x^2+1), o cuerpo Q(i) de los racionales complejos: $a+bi$, con a y b elementos de Q.

7. Clases residuales módulo x^2-2.—Ya que el polinomio x^2-2 no es descomponible en el cuerpo de los racionales, Q[x]/(x^2-2) es un cuerpo. Como en el ejemplo anterior, los restos son de primer grado y se obtienen reemplazando tantas veces como sea posible x^2 por 2. El cuerpo Q[x]/(x^2-2) es, por tanto, isomorfo al cuerpo de los racionales ampliado por la adjunción de un elemento ρ tal que $\rho^2=2$. Las reglas de cálculo son las siguientes:

adición: $$(a+b\rho)+(a'+b'\rho)=(a+a')+(b+b')\rho;$$

multiplicación:

$$(a+b\rho)(a'+b'\rho)=(aa'+2bb')+(ab'+ba')\rho;$$

división:

$$\frac{a'+b'\rho}{a+b\rho}=\frac{a'+b'\rho}{a+b\rho}\cdot\frac{a-b\rho}{a-b\rho}=\frac{(aa'-2bb')+(ab'-ba')\rho}{a^2-2b^2}$$

$a+b\rho$ y $a-b\rho$ son conjugados en Q(ρ). Su producto a^2-2b^2 es la norma de cada uno de ellos, ya que la norma del producto es igual al producto de las normas. Finalmente, el cambio de ρ en $-\rho$ es un automorfismo de Q(ρ), el único salvo el automorfismo neutro. En Q(ρ) el polinomio x^2-2 se descompone en

$$x^2-2=x^2-\rho^2=(x+\rho)(x-\rho).$$

Algebra y análisis.—Quizá el lector se habrá hecho la siguiente pregunta: «¿No es este ρ lo que siempre he llamado $\sqrt{2}$?».

Respuesta: La definición rigurosa del número $\sqrt{2}$ en cuanto número real irracional hace intervenir sucesiones *indefinidas* de raciona-

les. Se ha estudiado a propósito de los números reales, que, con sus propiedades especiales de naturaleza topológica (proximidad o entorno, continuidad, etc.), forman la base del *análisis* clásico. Hemos definido ρ por un proceso *finito* que pertenece por completo al dominio del *álgebra* en tanto que «ciencia de las cuatro operaciones».

Sin embargo, sobre la base del isomorfismo de $Q(\rho)$ y de $Q(\sqrt{2})$ es posible, si se plantea la necesidad de ello en alguna cuestión, identificar uno de los elementos ρ o $-\rho$ con $\sqrt{2}$.

8. Cuerpo de ruptura de un polinomio.—Sea $P(x)$ un polinomio no descomponible en $R[x]$. Construyamos el cuerpo $R[x]/P(x)$, isomorfo a la extensión $R(a)$ del cuerpo de los reales obtenida por adjunción de un elemento a tal que $P(a)=0$. En $R(a)$ el polinomio $P(x)$ se descompone: existe al menos la posibilidad de destacar $(x-a)$ como factor. Sean $P_1(x)$, $P_2(x)$, ... los factores irreducibles obtenidos. Se aplicará el mismo método haciendo la adjunción a $R(a)$ de un elemento b tal que $P_1(b)=0$, etc. Por la adjunción de un número finito de elementos $a, b, c, \ldots$ obtendremos un cuerpo $R(a, b, c, \ldots)$ en el que $P(x)$ se descompone por completo en factores de primer grado. Teniendo en cuenta el grado se llega a la siguiente conclusión:

TEOREMA.—*Para todo polinomio $P(x)$ de grado n y coeficientes reales, es posible formar por vía puramente algebraica y finita un cuerpo de ruptura en el que $P(x)$ se descompone por completo en n factores de primer grado. En este cuerpo la ecuación $P(x)=0$ admite n raíces.*

EJERCICIOS:

1. En el cuerpo $Q(i)$ se considera el conjunto $Z(i)$ de los elementos $a+bi$ con a y $b \in Z$ (enteros de Gauss). Demuéstrese que $Z(i)$ es un dominio de integridad. ¿Cuál es su cuerpo de cocientes?

2. Determínense los elementos unitarios del anillo $Z(i)$ definido en 1.

3. Demuéstrese que, en el cuerpo $Q(\sqrt{2})$, todo subcuerpo que contiene a todos los racionales es o bien Q, o bien todo el cuerpo $Q(\sqrt{2})$.

4. ¿Es un isomorfismo de anillo la correspondencia biunívoca

$$a+b\sqrt{2} \Longleftrightarrow a+b\sqrt{3}$$

entre los cuerpos $Q(\sqrt{2})$ y $Q(\sqrt{3})$?

5. Demuéstrese que, si se trata de establecer un isomorfismo entre $Q(\sqrt{2})$ y $Q(\sqrt{3})$, se vuelve a encontrar la ley de correspondencia examinada anteriormente. ¿Conclusión?

6. Estúdiese el cuerpo $Q[x]/(x^2-x-1)$ o, lo que es lo mismo, la extensión de Q por adjunción de un elemento u tal que $u^2=u+1$. Reglas de cálculo, automorfismos.

7. Estúdiese el anillo $K[x]/(x^2+1)$: *a)* en el caso $K=Z/5$; *b)* para $K=Z/7$. Se observará que en un caso x^2+1 es descomponible y en el otro no.

8. Estúdiese el anillo $Q[x]/x^2$: reglas de cálculo. Demuéstrese que es isomorfo al anillo de los números duales $a+b\epsilon$, con $\epsilon^2=0$. Determínense todos los divisores de cero.

9. Estúdiese el cuerpo finito $K[x]/(x^2-x-1)$ para $K=Z/3$. Constrúyanse las tablas de adición y de multiplicación empleando un u tal que $u^2=u+1$. (Un cuerpo finito es un campo de Galois.) [1]

DESCOMPOSICION EN LOS CUERPOS C Y R

9. Teorema de D'Alembert.—Hacia 1625, Girard afirmó que toda ecuación de grado n admite n raíces reales u «ocultas». Los científicos del siglo XVII emplearon con prudencia y desconfianza estos números que Descartes llamó «imaginarios». En el siglo XVIII se desarrollaron merced a la audacia y al virtuosismo de Euler. Fue Gauss quien primero *demostró* la proposición de Girard; dio de la misma varias demostraciones diferentes en el transcurso de su larga carrera. Su proposición se conoce en Francia con el nombre de teorema de D'Alembert y universalmente como el *teorema fundamental del álgebra*. De hecho, necesita utilizar recursos del análisis (no algebraicos).

TEOREMA.—*Todo polinomio de grado n y coeficientes complejos admite n raíces complejas (distintas o no).*

Para probarlo basta demostrar el siguiente teorema:

TEOREMA.—*Todo polinomio de grado n y coeficientes complejos admite al menos una raíz.*

La siguiente demostración reduce al mínimo la parte no algebraica.

Demostremos en primer lugar que se puede reducir al caso de un polinomio de coeficientes reales. En efecto, sea π un polinomio de coeficientes complejos, y $\bar{\pi}$, su conjugado; se sabe que el polinomio $P=\pi\bar{\pi}$ tiene sus coeficientes reales (véase pág. 219). Si admite una raíz $\xi=s+it$, se tiene o $\pi(s+it)=0$, o bien $\bar{\pi}(s+it)=0$; pero en este

[1] Señalemos que todo cuerpo finito es, necesariamente, conmutativo.

último caso, tomando conjugados se deduce que $\pi(s-it)=0$. La existencia de una raíz para P implica siempre la existencia de una raíz para π.

Consideremos el conjunto de los polinomios de coeficientes reales y dividámoslo en clases disjuntas: una de ellas comprenderá todos los polinomios cuyo grado es un entero divisible por cierta potencia de 2, p. ej., 2^m (y no 2^{m+1}). La primera clase incluye todos los polinomios de grado impar, y la segunda, todos los de grado divisible por 2 (y no por 4), etc. El teorema es cierto para la primera clase. En efecto, hemos visto que todo polinomio de coeficientes reales y grado impar admite al menos una raíz real.

Procedamos por recurrencia suponiéndolo cierto para la m-ésima clase (grado divisible por 2^{m-1}) y demostrando que también lo es para la clase siguiente. Sea, pues, A, de grado n, un polinomio de esta última clase. Sabemos que existe una extensión algebraica de R en la que A se descompone en n factores

$$A=(x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n).$$

Existen $\frac{1}{2}n(n-1)$ combinaciones de los α_j tomados de dos en dos. Elijamos $\omega=\frac{1}{2}n(n-1)+1$ números reales distintos, $a_1, a_2, \dots, a_\omega$, y para cada a_i (i de 1 a ω) formemos las $\frac{1}{2}n(n-1)$ expresiones de la forma

$$\alpha_j\alpha_k+a_i(\alpha_j+\alpha_k),$$

después el polinomio

$$p_i(x)=\prod_{j,k}[x-\alpha_j\alpha_k-a_i(\alpha_j+\alpha_k)],$$

en el que el producto del segundo miembro se extiende a todas las combinaciones de los α_j tomados de 2 en 2. Los coeficientes de p_i son polinomios simétricos de los α y, en consecuencia, reales[1]. Por otra parte, el grado de p_i es $\frac{1}{2}n(n-1)$ y p_i pertenece a la m-ésima clase; por la hipótesis de recurrencia existe una raíz en el cuerpo R(i), sea

$$\xi_i=\alpha_{j_i}\alpha_{k_i}+a_i(\alpha_{j_i}+\alpha_{k_i}) \quad \text{con} \quad \xi_i\in R(i).$$

Como hay $\omega=\dfrac{n(n-1)}{2}+1$ raíces del tipo ξ_i y $\dfrac{n(n-1)}{2}$ combinaciones de los α tomados de dos en dos, existirán forzosamente dos ξ

[1] Aplicando el teorema fundamental sobre los polinomios simétricos y las relaciones entre los coeficientes y raíces (véanse págs. 245 y 276).

formadas con la misma combinación de los índices. Cabe suponer los números atribuidos de tal suerte que esta sea

$$\xi_1 = \alpha_1\alpha_2 + a_1(\alpha_1 + \alpha_2) \quad \text{y} \quad \xi_2 = \alpha_1\alpha_2 + a_2(\alpha_1 + \alpha_2),$$

pero entonces

$$\alpha_1 + \alpha_2 = \frac{\xi_2 - \xi_1}{a_2 - a_1} = \lambda \quad \text{y} \quad \alpha_1\alpha_2 = \frac{a_2\xi_1 - a_1\xi_2}{a_2 - a_1} = \mu$$

pertenecen también a R(i). Por consiguiente, $\alpha_1 + \alpha_2$ y $\alpha_1\alpha_2$ figuran en R(i) como raíces de la ecuación de segundo grado $x^2 - 2\lambda x + \mu = 0$.

El teorema queda, pues, demostrado para la clase $(m+1)$-ésima.

10. Descomposición en C.—El teorema de D'Alembert se deduce fácilmente del teorema anterior. Sea P un polinomio de grado n de C[x]. Tenemos una raíz $z_1 \in C$, de donde

$$P = (x - z_1)P_1$$

y P_1 es de grado $n-1$. Aplicando de nuevo el teorema precedente, $P_1 = (x - z_2)P_2$, ... Finalmente, se obtiene en C[x] una descomposición completa

$$P = \lambda(x - z_1)(x - z_2) \ldots (x - z_n) \qquad (\lambda \in C). \qquad [1]$$

Por tanto, los únicos polinomios primos en el anillo C[x] son los binomios de primer grado. Las relaciones entre las raíces y los coeficientes (véase pág. 245) son siempre válidas en el cuerpo de los complejos.

La descomposición [1] puede contener factores idénticos. Si los agrupamos, se obtiene

$$P = \lambda(x - z_1)^{\alpha_1}(x - z_2)^{\alpha_2} \ldots (x - z_l)^{\alpha_l} \qquad [2]$$

con $\alpha_1 + \alpha_2 + \ldots + \alpha_l = n$. Se sabe que esta descomposición es posible solo de una manera. Se dice que z_1 es raíz múltiple de orden α_1, z_2 raíz de orden α_2, etc.

11. Descomposición sobre R.—El anillo R[x] es un subanillo de C[x] y, por consiguiente, todo polinomio P de coeficientes reales admite una descomposición completa en C[x]:

$$P = \lambda(x - z_1)^{\alpha_1} \ldots (x - z_l)^{\alpha_l} \qquad (\lambda \in R). \qquad [1]$$

TEOREMA.—*Las raíces en* C *de un polinomio de* R[*x*] *son o bien reales, o bien imaginarias conjugadas dos a dos, y con el mismo orden de multiplicidad.*

En efecto, como **P** pertenece a R[*x*], se confunde con el polinomio conjugado (véase pág. 219) obtenido tomando en [1] los conjugados de los coeficientes, de donde

$$\mathbf{P} = \lambda(x - \overline{z_1})^{\alpha_1}(x - \overline{z_2})^{\alpha_2} \dots (x - \overline{z_l})^{\alpha_l} \qquad [1 \text{ bis}]$$

Pero la unicidad de la descomposición en C[*x*] exige que [1] y [1 bis] coincidan. Por tanto, se tendrá:

o bien $z_i = \overline{z_i}$, de donde z_i real;

o bien $\overline{z_i} = z_j$, lo que exige $\alpha_i = \alpha'_j$ para que la descomposición sea única; las raíces son, por tanto, conjugadas.

En [1] se agrupan los factores complejos conjugados

$$(x - z_i)^{\alpha_i}(x - \overline{z_i})^{\alpha_i} = (x^2 + px + q)^{\alpha_i};$$

el trinomio $x^2 + px + q$ es de coeficientes reales y tal que $p^2 - 4q < 0$ (carece de raíces reales). Se deduce de ello la forma general de la descomposición en R[*x*],

$$\mathbf{P} = \lambda(x - a_1)^{\alpha_1}(x - a_2)^{\alpha_2} \dots (x - a_l)^{\alpha_l}(x^2 + p_1x + q_1)^{\beta_1} \dots (x^2 + p_kx + q_k)^{\beta_k}$$

y la propiedad expresada por el teorema siguiente:

TEOREMA.—*Los únicos polinomios primos en* R[*x*] *son los binomios de primer grado y los trinomios de segundo grado con discriminante negativo. Todo polinomio admite una descomposición única de la forma dicha.*

12. Ejemplo 1.º: ecuación de tercer grado.—La ecuación $x^3 + ax^2 + bx + c = 0$ se convierte en $y^3 + py + q = 0$ mediante la sustitución $y = x + a/3$. He aquí el artificio imaginado por Tartaglia (1506-1559) y expuesto por Geronimo Cardano (1501-1576). Hagamos $y = u + v$, con lo cual se obtiene $u^3 + v^3 + (u + v)(3uv + p) + q = 0$. Particularicemos u y v por $u^3 + v^3 + q = 0$, de donde $3uv + p = 0$. Tenemos:

$$u^3 + v^3 = -q, \qquad u^3v^3 = -\frac{p^3}{27},$$

y de aquí:

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}, \qquad v^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}};$$

si $\frac{q^2}{4} + \frac{p^3}{27} \geqslant 0$ se obtiene una raíz real (doble en el caso de la igualdad):

$$y = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

si $\frac{q^2}{4} + \frac{p^3}{27} < 0$ las tres raíces son reales, pero la fórmula de Cardano las da bajo forma aparentemente compleja: $y_1 = u + v$, $y_2 = ju + j^2 v$, $y_3 = j^2 u + jv$, designando u y v una de las raíces cúbicas de u^3 o de v^3. Los antiguos matemáticos llamaban a este el «casus irreducibilis». Leibniz veía en estas fórmulas imaginarias con resultado real una manifestación de la Providencia. Es necesario confesar que las fórmulas de Cardano son de singular estructura y plantean la cuestión de saber si otras fórmulas con radicales no darían cálculos reales para las raíces reales. Se demuestra que tales fórmulas no pueden existir.

Es posible dar, evidentemente, fórmulas trigonométricas, ya que la resolución de la ecuación de tercer grado se reduce a la trisección de un ángulo en el caso de raíces reales.

En efecto, se tiene $\cos 3\varphi = 4\cos^3\varphi - 3\cos\varphi$, y basta poner

$$y = r\cos\varphi, \qquad \text{con} \qquad r = 2\sqrt{-\frac{p}{3}} \qquad (\text{porque } p < 0)$$

para escribir $y^3 + py + q = 0$ en la forma

$$4\cos^3\varphi - 3\cos\varphi = \cos\psi = \frac{-q}{2\sqrt{-\frac{p^3}{27}}},$$

de donde

$$y_1 = r\cos\frac{\psi}{3}, \qquad y_2 = r\cos\frac{2\pi + \psi}{3}, \qquad y_3 = r\cos\frac{4\pi + \psi}{3}$$

13. Ejemplo 2.°: ecuación de cuarto grado.—La ecuación

$$x^4 + ax^3 + bx^2 + cx + d = 0.$$

se reduce a la forma normal $y^4 + py^2 + qy + r = 0$ haciendo $y = x + a/4$. Tomamos como incógnitas auxiliares ξ, η, θ tales que

$$x^4 + px^2 + qx + r = (x^2 + \xi x + \eta)(x^2 - \xi x + \theta);$$

se obtienen las condiciones necesarias y suficientes

$$\eta+\theta=\xi^2=p, \qquad \xi(\theta-\eta)=q, \qquad \eta\theta=r.$$

De las dos primeras, se deduce que

$$\eta=\frac{1}{2}\left(p+\xi^2-\frac{q}{\xi}\right), \qquad \theta=\frac{1}{2}\left(p+\xi^2+\frac{q}{\xi}\right)$$

y sustituyendo en la tercera y poniendo $\xi^2=\zeta$,

$$\zeta^3+2p\zeta^2+(p^2-4r)\zeta-q^2=0 \qquad \text{(ecuación resolvente).}$$

Esta ecuación de tercer grado tiene, en general, tres raíces distintas, ζ_1, ζ_2, ζ_3; designemos por ξ_1 (resp. ξ_2) una de las raíces cuadradas de ζ_1 (resp. ζ_2), a la que corresponderán por un cálculo racional η_1 y θ_1 (resp. η_2 y θ_2). Como en general, $\xi_1\neq\pm\xi_2$, tendremos dos descomposiciones distintas:

$$(x^2+\xi_1x+\eta_1)(x^2-\xi_1x+\theta_1)=(x^2+\xi_2x+\eta_2)(x^2-\xi_2x+\theta_2),$$

de donde,

$$\begin{aligned}
x^2+\xi_1x+\eta_1&=(x-\alpha_1)(x-\alpha_2),\\
x^2-\xi_1x+\theta_1&=(x-\alpha_3)(x-\alpha_4),\\
x^2+\xi_2x+\eta_2&=(x-\alpha_1)(x-\alpha_3),\\
x^2-\xi_2x+\theta_2&=(x-\alpha_2)(x-\alpha_4);
\end{aligned}$$

se ve, pues, que $x-\alpha_1$, m. c. d. del primero y tercer trinomios, se calcula racionalmente en función de ξ_1, η_1, ξ_2, η_2, etc.

En resumen, la ecuación de cuarto grado puede resolverse por medio de una resolvente de tercer grado y la extracción de dos raíces cuadradas.

Se observa que por medio de una conveniente ordenación es posible siempre obtener $\xi_1=\alpha_1+\alpha_2$; por otra parte, como en la ecuación no hay términos en x^3, tenemos $\alpha_3+\alpha_4=-(\alpha_1+\alpha_2)$, de donde

$$\xi_1=\frac{1}{4}(\alpha_1+\alpha_2-\alpha_3-\alpha_4) \qquad \text{y} \qquad \zeta_1=\xi_1^2=\frac{1}{4}(\alpha_1+\alpha_2-\alpha_3-\alpha_4)^2.$$

Esto explica por qué ζ_1 satisface a una ecuación de tercer grado: aplicando a $\alpha_1+\alpha_2-\alpha_3-\alpha_4$ todas las sustituciones de S_3 se obtienen solo tres valores, y el polinomio

$$(\zeta-\zeta_1)(\zeta-\zeta_2)(\zeta-\zeta_3)$$

tiene sus coeficientes en el cuerpo base.

EJERCICIOS:

1. Resuélvanse las ecuaciones

$$x^3+6x+2=0 \qquad \text{y} \qquad x^3-3abx+a^3+b^3=0.$$

2. En C[x] descompóngase

$$P=2x^3-(5+6i)x^2+9ix+1-3i,$$

sabiendo que admite una raíz real.

3. Descompóngase $P=(x+i)^n-(x-i)^n$ en producto de factores.

4. Descompóngase $P=x^6-i$ por un método algebraico. Dedúzcase de ello las razones trigonométricas del arco $\pi/12$.

5. Resuélvanse las siguientes ecuaciones:

$$x^4-2x^3+4x^2-2x+3=0; \qquad x^4-x^3-3x^2+x+1=0.$$

6. Sea la ecuación $x^3+x+a=0$, en la que a es real. Hállese a de manera que el valor absoluto de la raíz real sea mayor que el módulo común de las dos raíces complejas.

7. ¿Cómo hay que elegir p y q para que $x^3+px+q=0$ admita como raíces los afijos de los vértices de un triángulo isósceles? Por un cambio de incógnita, $x=\xi u$ (donde u es un parámetro a determinar), redúzcase al caso en que dos raíces sean complejos conjugados, y la otra, real. Se halla una condición para la razón p^3/q^2.

8. ¿Cómo hay que elegir p, q, r para que las raíces de $x^4+px^2+qx+r=0$ admitan como imágenes los vértices de un paralelogramo, o de un rectángulo, o de un rombo, o de un cuadrado?

9. Descompóngase el polinomio $(1-x^2)^3+8x^3$ en un producto de factores reales.

10. Hállese la condición para que la ecuación

$$x^4+x^3+(a+ib)x^2+(c+id)x+p+iq=0$$

admita dos raíces complejas conjugadas.

11. Efectuando la división del polinomio $x^4+14x^3+50x^2+7x-10$ por x^2+7x+1, de la forma del cociente dedúzcase la descomposición en factores irreducibles en R[x] del polinomio dado.

12. Exprésese que $P=x^3+px+q$ tiene una raíz real y dos raíces complejas de módulo 2. Resuélvase la ecuación en este caso.

DERIVADA DE UN POLINOMIO. RAICES MULTIPLES

En álgebra elemental se estudian los polinomios en el cuerpo R definido empíricamente y se introduce la derivada de un polinomio A como el límite del cociente de incrementos $[A(x+\Delta x)-A(x)]/\Delta x$, cuan-

do $\Delta x = 0$. Esta definición, que hace intervenir el orden de R y la continuidad, no podría generalizarse a polinomios sobre un cuerpo cualquiera. Conviene, pues, eliminar las hipótesis inútiles y conservar únicamente el núcleo algebraico.

Sin necesidad de introducir un rigorismo perfecto en las consideraciones que siguen, a las que solo pedimos que sean sugerentes, decimos que la definición de derivada por

$$A' = \lim_{\Delta x \to 0} [A(x + \Delta x) - A(x)]/\Delta x$$

implica que para Δx finito y muy pequeño la cantidad $A(x + \Delta x) - A(x)$ difiere de la $A' \cdot \Delta x$ en un término correctivo que es al menos del orden de $(\Delta x)^2$. Este término pequeño, de segundo orden, que el físico desprecia en sus cálculos, el matemático puede identificarlo a cero si se coloca en el punto de vista que sigue.

14. Definición algebraica formal de la derivada.—A partir del polinomio $A(x) \in K[x]$, formemos el polinomio $A(x + y) \in K[x, y]$, en el que y generaliza el papel demasiado especializado de Δx. Para que y^2 no sea únicamente «despreciable», sino rigurosamente reemplazable por cero, basta considerar las congruencias módulo y^2, lo que, según el estudio general del número 4 (pág. 246), nos conduce a considerar el anillo $K[x][y]/y^2$. Toda clase residual admite un representante de la forma $A(x) + yB(x)$.

Las observaciones anteriores nos llevan, pues, a establecer como *definición de la derivada* A':

$$A(x + y) - A(x) \equiv y \cdot A'(x) \qquad (\text{mód. } y^2).$$

Esta relación define $A'(x)$ de manera única.

Demostremos que se conservan las reglas usuales para la suma y el producto. Para la suma esto es evidente:

Si $$S = A + B,$$

$$S(x + y) - S(x) = [A(x + y) - A(x)] + [B(x + y) - B(x)],$$

de donde

$$S' = A' + B'.$$

Para el producto, si $P = A \cdot B$ con $A(x+y) \equiv A(x) + yA'(x)$ y $B(x+y) \equiv$ $\equiv B(x) + yB'(x)$, se tiene:

$$P \equiv A(x)B(x) + y[A(x)B'(x) + A'(x)B(x)],$$

de donde

$$P'(x) = A(x)B'(x) + A'(x)B(x).$$

Se obtienen así las fórmulas clásicas. En particular, para $A(x) = x^m$,

$$(x+y)^m \equiv x^m + mx^{m-1}y \qquad (\text{mód. } y^2),$$

de donde

$$A'(x) = mx^{m-1}.$$

Señalemos, sin embargo, que, en un cuerpo de característica p, la derivada de x^p es px^{p-1}, polinomio que toma el valor cero para todo $x \in K$ y lo mismo para todo x elemento de un supercuerpo de K. En ello hay una fuente de dificultades para determinadas teorías. En lo que sigue nos limitaremos al caso en el que la característica es infinita (salvo mención expresa de lo contrario).

EJERCICIOS:

1. Hállense de dos maneras distintas las derivadas de los polinomios

$$(1x-2)^3 \quad \text{y} \quad (1x^2-2x+1)(1x^2+1x-2) \quad \text{en} \quad Z/5[x].$$

2. Calcúlese la derivada con relación a x del polinomio $A(ax+b)$ [obtenido reemplazando u por $ax+b$ en el polinomio $A(u)$].

3. Con mayor generalidad, calcúlese la derivada de $A[B(x)]$.

15. Derivadas sucesivas.—Definida la derivada, o derivada primera de un polinomio $A(x)$, se define la derivada segunda como derivada de la derivada primera, etc. Notaciones:

$$A', A'', A''', A^{IV} \quad \text{o} \quad A^{(4)}, \ldots, A^{(n)}.$$

Las derivadas sucesivas del polinomio x^m son, por consiguiente, $A' = mx^{m-1}$, $A'' = m(m-1)x^{m-2}$, ..., $A^{(p)} = m(m-1) \ldots (m-p+1)x^{m-p}$ para $m \leqslant p$, $A^{(m)} = m(m-1) \ldots 3.2.1 = m!$ y $A^{(q)} = 0$ para $q > m$.

16. Fórmula de Taylor.—Sea el polinomio

$$B = b_0 + b_1x + b_2x^2 + \ldots + b_mx^m.$$

Escribimos las igualdades

$$\begin{array}{l|l} b_0 & 1=1 \\ b_1 & a+h=a+h \\ b_2 & (a+h)^2=a^2+2ah+h^2 \\ b_3 & (a+h)^3=a^3+3a^2h+3ah^2+h^3 \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_m & (a+h)^m=a^m+ma^{m-1}h+\dfrac{m(m-1)}{2}a^{m-2}h^2+ \\ & \qquad +\dfrac{m(m-1)(m-2)}{3}a^{m-3}h^3+\ldots+h^m \end{array}$$

Multipliquemos los dos miembros de las igualdades que figuran en la primera, segunda, ... filas por b_0, b_1, ..., respectivamente, y sumemos miembro a miembro, teniendo en cuenta las columnas; se obtiene:

$$B(a+h)=B(a)+hB'(a)+\frac{h^2}{2!}B''(a)+\frac{h^3}{3!}B'''(a)+\ldots+\frac{h^m}{m!}B^{(m)}(a),$$

que es la fórmula de Taylor, la cual puede escribirse también, haciendo $h=x-a$,

$$B(x)=B(a)+(x-a)B'(a)+\frac{(x-a)^2}{2!}B''(a)+ \\ +\frac{(x-a)^3}{3!}B'''(a)+\ldots+\frac{(x-a)^m}{m!}B^{(m)}(a).$$

En particular, para $a=0$, se tiene:

$$B(x)=B(0)+xB'(0)+\frac{x^2}{2!}B''(0)+\ldots+\frac{x^m}{m!}B^{(m)}(0).$$

Se sabe, pues, escribir un polinomio de grado m en cuanto se conocen los valores de él y de sus m primeras derivadas para $x=0$.

17. Raíces múltiples.—El elemento a es una raíz múltiple de orden λ del polinomio $B(x)$ si se tiene que $B(x)=(x-a)^\lambda C(x)$ con $C(a)\neq 0$.

TEOREMA.—*Para que a sea una raíz múltiple de orden λ del polinomio $B(x)$ es necesario y suficiente que*

$$B(a)=B'(a)=\ldots=B^{(\lambda-1)}(a)=0 \qquad \text{con} \qquad B^{(\lambda)}(a)\neq 0.$$

Es una consecuencia inmediata de la fórmula de Taylor en su segunda forma.

Complemento.—Sea el polinomio $A = P^m B$, donde P es primo con B. Tenemos

$$A' = P^{m-1}(PB' + mBP') = P^{m-1}Q,$$

siendo P todavía primo con Q; por consiguiente, un factor múltiple irreducible y de orden m para un polinomio A es un factor múltiple de orden $m-1$ para su derivada.

De ello resulta que si $A = P_1^{\alpha_1} P_2^{\alpha_2} \dots P_r^{\alpha_r}$, donde los factores P_i son irreducibles, el producto $A^* = P_1 P_2 \dots P_r$ se obtiene sin otros cálculos que el de A' y el del M.C.D., Δ, de A y de A': basta observar que $A^* \Delta = A$.

TEOREMA.—*Si* A(x) *carece de factores múltiples en el cuerpo* K, *tampoco los tiene en una extensión de* K.

El hecho de que los factores de A sean simples implica que A y A' son primos entre sí, de donde la relación de Bezout $1 = SA + TA'$. Esta igualdad válida en K será también válida en una extensión del cuerpo y, por consiguiente, A y A' continúan siendo primos entre sí.

En particular, un polinomio irreducible carece de raíces múltiples en cualquier extensión del cuerpo base (véase ejercicio 12-6, pág. 269).

EJERCICIOS:

1. Sobre un cuerpo cualquiera de característica infinita, determínense α y β para que el polinomio $\alpha x^{n+1} + \beta x^n + 1$ sea divisible por $(x-1)^2$.

2. Estúdiese si el polinomio $x^5 - 5x - a$ puede tener raíces múltiples sobre Q, sobre Q(i).

3. Si P(x) designa un polinomio cualquiera, demuéstrese que la ecuación

$$\frac{x-a}{2}[P'(x) + P'(a)] + P(a) - P(x) = 0$$

admite $x = a$ como raíz múltiple. ¿Cuál es su orden?

4. Todo polinomio de cuarto grado cuyas derivadas primera y tercera se anulan para un mismo valor de la variable es bicuadrado (cuadrado de un polinomio de segundo grado).

5. Si p es un polinomio de grado n, demuéstrese que P y sus n primeras derivadas son linealmente independientes.

EL CUERPO DE LAS FRACCIONES RACIONALES

18. Cuerpo de las fracciones del anillo K[x].—La división exacta es, en general, imposible sobre K[x]: si A y B son dos polinomios, sólo es posible encontrar un tercer polinomio Q que satisfaga a la igualdad $A = BQ$ cuando B es un divisor de A. Para que esta operación sea siempre posible nos vemos obligados a construir el *cuerpo de las fracciones* del anillo K[x] (anillo de integridad). Se considera el conjunto de pares (A, B), dotado de la relación de equivalencia $(A, B) \equiv (A', B')$ si, y solo si, $AB' = BA'$.

Es fácil demostrar que el conjunto cociente adquiere la estructura de cuerpo cuando se definen sobre los pares las dos operaciones

$$(A, B) + (C, D) = (AD + BC, BD),$$
$$(A, B) \cdot (C, D) = (AC, BD),$$

estables respecto a la relación de equivalencia.

Carece de objeto desarrollar la teoría; el razonamiento utilizado para construir los racionales es válido para todo anillo de integridad y proporciona su cuerpo de cocientes.

En el cuerpo obtenido, el subconjunto (A, *e*), donde *e* es el elemento unidad de K, es isomorfo a K[x]. La identificación habitual hace que un elemento cualquiera (A, B) del cuerpo de las fracciones aparezca como cociente de los elementos A y B, de donde la notación corriente A/B. De ello resulta:

TEOREMA.—*El anillo* K[x] *puede, salvo un isomorfismo, sumergirse en un cuerpo en el que todo elemento es el cociente de dos polinomios del anillo inicial.*

Hemos obtenido de esta forma el cuerpo de las fracciones racionales, denotado por K(x).

La igualdad de dos fracciones A/B y C/D viene definida por la relación de equivalencia, $AD = BC$, que conduce, cuando los polinomios A y B no son primos entre sí, a la simplificación de una fracción. Las operaciones sobre las fracciones racionales se han indicado a propósito del estudio de las reglas de cálculo sobre un cuerpo.

Parte entera.—Dada la fracción A/B, efectuemos el cociente entero de A por B:

$$A = BQ + A' \qquad \text{de donde} \qquad \frac{A}{B} = Q + \frac{A'}{B}$$

Puesto que el grado de A' es inferior al de B, se llega al resultado que sigue: toda fracción racional puede escribirse en forma de polinomio, más una fracción en la que el grado del numerador es inferior al del denominador. El polinomio se llama parte entera, y es nula cuando el grado de A es inferior al de B.

19. Descomposición en elementos simples.- Estudiemos en primer lugar dos teoremas preliminares:

TEOREMA 1.—*Toda fracción racional* A/B, *cuyo denominador sea descomponible en un producto* B_1B_2 *de dos factores primos entre sí, es la suma de dos fracciones que tienen por denominadores* B_1 *y* B_2, *respectivamente.*

En virtud de la igualdad de Bezout, existen dos polinomios S y T, tales que

$$1 = SB_1 + TB_2.$$

Multipliquemos por A:

$$A = ASB_1 + ATB_2.$$

Dividamos por B_1B_2:

$$\frac{A}{B} = \frac{AS}{B_2} + \frac{AT}{B_1}$$

OBSERVACIÓN.—Supongamos que grado de A $<$ grado de B y pongamos de manifiesto las partes enteras Q_1 y Q_2 de las fracciones obtenidas. Resulta así:

$$\frac{A}{B} = (Q_1 + Q_2) + \frac{\alpha}{B_1} + \frac{\beta}{B_2},$$

igualdad que es imposible, salvo para $Q_1 + Q_2 = 0$, dado la hipótesis hecha acerca de los grados. Demostremos que la descomposición obtenida de esta forma es única. Si existiera una segunda

$$\frac{A}{B} = \frac{\alpha'}{B_1} + \frac{\beta'}{B_2},$$

se deduciría la igualdad imposible $B_2(\alpha - \alpha') = B_1(\beta' - \beta)$, ya que B_1 (primo con B_2) debería dividir a $\alpha - \alpha'$, que es de menor grado.

TEOREMA 2.—*Toda función racional,* A/B, *cuyo denominador sea una potencia perfecta,* $B=[B_1(x)]^n$, *es la suma de un polinomio y de fracciones que tienen por denominadores las diversas potencias de* B_1 *inferiores o iguales a n, y por numeradores, polinomios de grados inferiores al de* B_1.

Dividimos A por B_1, de donde

$$A = B_1Q_1 + A_1 \qquad \text{con} \qquad \text{grado de } A_1 < \text{grado de } B_1.$$

En consecuencia,

$$\frac{A}{B} = \frac{Q_1}{B_1^{n-1}} + \frac{A_1}{B_1^n}$$

Hemos puesto así en evidencia una de las fracciones del enunciado, y el problema queda reducido al análogo correspondiente a $(B_1)^{n-1}$, por lo que basta reiterar el proceso: el polinomio es el cociente de la última división, que, por otra parte, coincide con la parte entera de la fracción dada. Se llega de este modo a la fórmula:

$$\frac{A}{B} = \frac{A_1}{(B_1)^n} + \frac{A_2}{(B_1)^{n-1}} + \ldots + \frac{A_n}{B_1} + E$$

y se demuestra fácilmente la unicidad de la descomposición obtenida.

Caso general.—Con una fracción cualquiera, el denominador puede escribirse en la forma

$$B(x) = [p_1(x)]^{n_1}\,[p_2(x)]^{n_2} \ldots [p_i(x)]^{n_i}.$$

Siendo los polinomios p_1, p_2, ..., p_i irreducibles y, por consiguiente, primos entre sí, lo mismo sucede con $p_1^{n_1}$, $p_2^{n_2}$, ...

En la fracción dada A/B separemos la parte entera:

$$\frac{A}{B} = E(x) + \frac{A'}{B},$$

y apliquemos el teorema 1:

$$\frac{A}{B} = E + \frac{A_1}{p_1^{n_1}} + \frac{A_2}{p_2^{n_2}} + \ldots + \frac{A_i}{p_i^{n_i}}, \qquad \text{con grado de } A_i < \text{grado de } p_i^{n_i}.$$

A cada fracción del segundo miembro, se aplica el teorema 2[1]:

$$\frac{A}{B} = E + \frac{A_1^1}{p_1^{n_1}} + \frac{A_1^2}{p_1^{n_1-1}} + \ldots + \frac{A_1^{n_1}}{p_1}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ \frac{A_l^1}{p_l^{n_l}} + \frac{A_l^2}{p_l^{n_l-1}} + \ldots + \frac{A_l^{n_l}}{p_l}$$

Los polinomios A_i^j de una misma línea tienen todos grados inferiores al del polinomio p_i correspondiente. Se deduce de ello fácilmente la unicidad de la descomposición.

Resumen.—Sobre un cuerpo K toda fracción racional puede expresarse como suma de un polinomio y de fracciones de la forma $q(x)/[p(x)]^m$, donde $p(x)$ es irreducible y $q(x)$ tiene grado inferior al de $p(x)$. Los denominadores $[p(x)]^m$ provienen de la descomposición en factores del denominador de la fracción dada: si, en esta descomposición, $p(x)$ figura con exponente n, tenemos $m \leqslant n$. Se dice que se ha descompuesto la fracción racional en *elementos* o *fracciones simples.*

20. Caso particular importante.—Sobre el cuerpo de los reales hemos visto que los únicos polinomios primos son los de primer grado y aquellos de segundo grado, $x^2 + px + q$, para los que $p^2 - 4q < 0$. La descomposición del denominador es, por consiguiente,

$$B = (x-a)^\alpha (x-b)^\beta \ldots (x^2 + p_1 x + q_1)^\lambda \ldots$$

Para un denominador $(x-a)^\alpha$, los numeradores serán constantes, y binomios de primer grado para $(x^2+px+q)^\lambda$. De donde, la descomposición

$$\frac{A}{B} = E + \sum_1^\alpha \frac{a_i}{(x-a)^{\alpha_i}} + \sum_1^\beta \frac{b_i}{(x-b)^{\beta_i}} + \ldots +$$

$$+ \sum_1^\lambda \frac{\mu_i x + \nu_i}{(x^2+px+q)^{\lambda_i}} + \ldots$$

[1] En la notación A_i^j el superíndice caracteriza el orden en una sucesión de polinomios; es decir, no es un exponente.

EJEMPLO.—Sobre R, descompóngase la fracción

$$\frac{A}{B} = \frac{x^7+5}{(x^2+x+1)^2(x+2)^3}.$$

El polinomio x^2+x+1 es primo sobre R y, por tanto,

$$\frac{A}{B} = \frac{\alpha x+\beta}{(x^2+x+1)^2} + \frac{\gamma x+\delta}{x^2+x+1} + \frac{\epsilon}{(x+2)^3} + \frac{\zeta}{(x+2)^2} + \frac{\eta}{(x+2)} + E.$$

La parte entera es igual al cociente de la división de A por B, y vale 1. La igualdad de los dos miembros implica (y recíprocamente)

$$\begin{aligned} x^7+5 = {} & (\alpha x+\beta)(x+2)^3 + (\gamma x+\delta)(x^2+x+1)(x+2)^3 + \\ & + \epsilon(x^2+x+1)^2 + \zeta(x+2)(x^2+x+1)^2 + \\ & + \eta(x^2+x+1)^2(x+2)^2 + (x^2+x+1)^2(x+2)^3. \end{aligned}$$

Para hallar el valor de ϵ se reemplaza x por -2, de donde

$$-123 = 9\epsilon, \qquad \epsilon = -\frac{41}{3}$$

Para obtener α y β se escribe que el resto de la división por x^2+x+1 es el mismo en ambos miembros.

A la izquierda,

$$x^7+5 = x(x^6-1)+x+5;$$

el resto es, pues, $x+5$, ya que

$$x^6-1 = (x^3+1)(x^3-1) = (x^3+1)(x-1)(x^2+x+1).$$

A la derecha,

$$(x+2)^3 = x^3+6x^2+12x+8 = (x^2+x+1)(x+5)+6x+3;$$

basta, pues, buscar el resto de $(\alpha x+\beta)(6x+3)$, es decir, de $6\alpha x^2 + (6\beta+3\alpha)x+3\beta$. Se obtiene restando $6\alpha x^2+6\alpha x+6\alpha$, de donde

$$(6\beta-3\alpha)x+3\beta-6\alpha = x+5,$$

$$\begin{cases} -3\alpha+6\beta = 1, \\ -6\alpha+3\beta = 5, \end{cases}$$

y, finalmente,

$$\alpha = -1, \qquad \beta = -\frac{1}{3}$$

Para facilitar la identificación, hagamos pasar en primer lugar a la izquierda el primero y el tercer términos de la suma, después dividiremos los dos miembros por $(x+2)(x^2+x+1)$.

A la izquierda, se obtiene

$$x^7+5+\left(x+\frac{1}{3}\right)(x+2)^3+\frac{41}{3}(x^2+x+1)^2=$$

$$=\frac{1}{3}(3x^7+44x^4+101x^3+165x^2+118x+64);$$

y dividiendo por

$$(x+2)(x^2+x+1)=x^3+3x^2+3x+2,$$

resulta a la izquierda

$$x^4-3x^3+6x^2+\frac{11}{3}x+\frac{32}{3}$$

y, a la derecha,

$$(\gamma x+\delta)(x+2)+\zeta(x^2+x+1)+\eta(x^2+x+1)(x+2)+(x^2+x+1)(x+2)^2.$$

Si se hace $x=-2$, se obtiene $\zeta=202/9$.

Identificando los restos de la división por x^2+x+1, resulta:

$$-\frac{4}{3}x+\frac{5}{3}=3\delta x+3\delta-3\gamma,$$

de donde,

$$\delta=-\frac{4}{9}, \qquad \gamma=-1$$

y se termina reemplazando x por -1, de donde $\eta=-7$; como comprobación se hace después $x=0$.

EJERCICIOS:

1. Justifíquese con detalle la afirmación del texto referente a la unicidad de la descomposición dada por el teorema 2.

2. Aplíquese el teorema 2 a la fracción $\dfrac{3x^2+2x-5}{(x^2-3x+2)^3}$.

3. Sobre el cuerpo Q, descompónganse en elementos simples las fracciones

$$\frac{3x^3+4}{(x+1)(x^2-2x+3)^2} \qquad \frac{2x^2-1}{(x^2+x+1)(x^2+1)^2} \qquad \frac{1}{x^2(1+x^3)}$$

4. Descompóngase en elementos simples la fracción $\frac{x^{2n}}{(x^2+1)^n}$ utilizando para ello el cambio de variable $y=x^2+1$.

5. Sobre un cuerpo cualquiera, ¿qué relación debe existir entre los coeficientes a, b, p, q para obtener la descomposición

$$\frac{x^2+px+q}{(x-a)^2(x-b)^2}=\frac{\lambda}{(x-a)^2}+\frac{\mu}{(x-b)^2}?$$

6. Sobre el cuerpo Z/7, descompóngase la fracción

$$\frac{2x^5+3x^3+5x^2-4}{(x^2+1)^2}$$

Trátese a continuación el problema sobre el cuerpo Z/5.

7. Sobre el cuerpo de los racionales descompóngase $\frac{x}{x^8-1}$ en fracciones simples. Trátese a continuación el mismo problema sobre el cuerpo de los reales.

EJERCICIOS SOBRE EL CAPITULO XII

12-1. Indíquese cómo se puede deducir de un polinomio dado otro polinomio congruente módulo x^2-x+1 y de grado inferior.

Dedúzcase de lo anterior cómo formar el resto de la división de un polinomio por x^2-x+1.

Demuéstrese que el polinomio $(x-1)^{n+2}+x^{2n+1}$ es divisible por x^2-x+1. Fórmese el cociente para $n=2$ y $n=3$.

12-2. Pruébese que el cuerpo $Q(\rho)$ (véase pág. 249) puede ordenarse por las condiciones siguientes: siendo $\rho>0$, se define $a+b\rho>0$ para

$$a>0, \quad b>0; \qquad a>0, \quad b<0$$

y

$$a^2>2b^2; \qquad a<0, \quad b>0 \quad \text{y} \quad a^2<2b^2.$$

Compruébense los postulados de orden en un cuerpo. Explíquense los convenios precedentes utilizando para ello el isomorfismo de $Q(\rho)$ y de $Q(\sqrt{2})$.

12-3. Sea $P(x)$ un polinomio del anillo $Q[x]$. Si la variable x toma sus valores en un supercuerpo $Q(i)$ o también $Q(\sqrt{N})$ (con N entero no cuadrado perfecto), el polinomio $P(x)$ llega a ser también un elemento del supercuerpo. Demuéstrese que, para dos valores conjugados $x=a+bi$ y $\bar{x}=a-bi$ (o también $x=a+b\sqrt{N}$, $\bar{x}=a-b\sqrt{N}$), el polinomio toma valores conjugados. Dedúzcase de ello que si $a+bi$ (o $a+b\sqrt{N}$) es raíz de orden α del polinomio, lo mismo sucede para $a-bi$ (o $a-b\sqrt{N}$).

12-4. Determínense los números racionales λ y μ de manera que la ecuación

$$x^4+4x^3-4\lambda x+4\mu=0$$

tenga una raíz doble de la forma $\alpha+\beta\sqrt{3}$, donde α y β son racionales. *(Resp.:* $\lambda=2$, $\mu=1$.)

12-4 *bis*. Reconózcase si la ecuación $x^3-5x-2=0$ tiene una raíz de la forma $a+b\sqrt{N}$.

12-5. Descompónganse en elementos simples sobre $Q(x)$ las fracciones racionales

$$\frac{2x^3+5x^2}{(x^2+1)^2(x-3)}, \qquad \frac{5x^4-4x+1}{(x^2-2)^2(x+1)^2},$$

utilizando para ello la siguiente observación. Como la descomposición resulta de un cálculo formal, la variable x puede tomar sus valores en un supercuerpo $Q(i)$ o $Q(\sqrt{N})$ del cuerpo base: los dos miembros se mantienen iguales.

Después de haber hecho enteros los dos miembros de la igualdad buscada, se elegirán valores sencillos de $Q(i)$ o de $Q(\sqrt{2})$ que permitan una determinación inmediata de algunos coeficientes.

12-6. Siendo K el cuerpo Z/3, se considera el cuerpo I de las fracciones racionales de una variable u con coeficientes en K, después el anillo L[x] y, finalmente, el polinomio x^3-2u. Extiéndase L por adjunción de un elemento φ tal que $\varphi^3-2u=0$ y conclúyase de ello que un polinomio irreducible puede conducir a raíces múltiples si la característica no es infinita (véase Sec. 17, Complemento).

12-7. La descomposición de un polinomio cualquiera $A(x)$ puede escribirse en la forma

$$A=(P_{11}\cdot P_{12}\dots P_{1\lambda})^i(P_{21}\cdot P_{22}\dots P_{2\mu})^j\dots=R_1^iR_2^j\dots,$$

donde los P_{ij} son irreducibles y todos diferentes. Demuéstrese que es posible obtener los R por división. Se utilizará el M. C. D. Δ_m de $(A^*)^m$ y de A, donde A^* se ha definido en la sección 17.

12-8. 1.º Sea un polinomio $A(x)$ del anillo $Q(x)$. Pruébese que existe un polinomio $P(x)$ y uno solo tal que

$$P(x+1)+P(x-1)=2A(x).$$

Grado de P. Demuéstrese que su parte par y su parte impar dependen, respectivamente, de las partes par e impar de $A(x)$.

2.º Suponemos $A(x)=x^n$. Se llama P_n al polinomio P correspondiente (uno para cada valor de n). Demuéstrese que

$$P_n'(x)=nP_{n-1}(x).$$

Determínese el polinomio P_n.

12-9. En todo el problema nos referimos a polinomios en x cuyos coeficientes son números racionales.

1.º Dado un polinomio $g(x)$, demuéstrese que existe un polinomio $f(x)$ y uno solo que verifica las condiciones

$$f(0)=0, \qquad f(x)-f(x-1)=g(x).$$

2.º Se designa por $f_p(x)$ el polinomio obtenido resolviendo la cuestión anterior cuando se toma $g(x)=x^p$ (p entero $\geqslant 0$; por definición, el polinomio x^0 es la constante 1).

Pruébese que $f_p(x)$ es divisible por $x+1$ para $p\geqslant 1$.

Establézcase, para todo entero positivo n, la relación

$$f_p(n)=1^p+2^p+\ldots+n^p.$$

3.º Se designa por $f'_p(x)$ la derivada del polinomio $f_p(x)$. Demuéstrese que

$$f'_p(x)=pf_{p-1}(x)+f'_p(0) \qquad \text{para} \qquad p\geqslant 1.$$

4.º Sea $h_p(x)$ el polinomio que tiene por derivada $f_p(x)$ y que se anula para $x=0$. Dedúzcase la relación

$$f_{p+1}(x)=(p+1)h_p(x)-[(p+1)h_p(1)-1]x.$$

Calcúlese explícitamente $f_p(x)$ para $p=1, 2, 3, 4$.

5.º Se consideran los tres polinomios

$$P(x)=1+x+x^2+\ldots+x^n,$$
$$Q(x)=1+2x+3x^2+\ldots+(n+1)x^n.$$
$$R(x)=1+2^2x+3^2x^2+\ldots+(n+1)^2x^n.$$

Calcúlese el coeficiente de x^n en el desarrollo del polinomio producto $P(x)Q(x)R(x)$. Se escribirá el coeficiente en la forma más sencilla posible.

(*M.G.P.*, Poitiers.)

CAPITULO XIII

POLINOMIOS DE VARIAS VARIABLES

ANILLO K[x, y]

1. Polinomios formales.—La teoría de los polinomios formales se extiende al caso de varias variables. Dado que las fórmulas se complican rápidamente estudiamos solo el caso de dos variables, pero el lector se dará cuenta de que el método es general.

Se llama polinomio formal de dos variables, sobre un cuerpo K, toda sucesión de dos índices a_{ij} (i y j enteros positivos o nulos) de elementos de K que no contiene más que un número finito de términos no nulos. Los términos que corresponden a $i+j=p$ se llaman de grado p. Se escribe la sucesión ordenándola según los grados crecientes y, dentro de un mismo grado, según los índices i decrecientes. Tenemos, pues, para un polinomio de grado n:

$$P=[a_{00}, a_{10}, a_{01}, a_{20}, a_{11}, a_{02}, a_{30}, \ldots, a_{0n}, 0, 0, \ldots].$$

En general, escribimos $P=[a_{ij}]$. Una sucesión que contiene solo un término no nulo es un monomio, y está caracterizado por su coeficiente y por sus dos índices.

2. Espacio vectorial.—Sean P y Q dos polinomios dados a los que se asocia el polinomio $P+Q$ obtenido sumando los términos del mismo lugar:

$$P=[a_{ij}], \qquad Q=[b_{ij}], \qquad P+Q=[a_{ij}+b_{ij}].$$

Esta operación dota al conjunto de estructura de grupo aditivo. El elemento neutro es el polinomio nulo, O, cuyos coeficientes son todos nulos (no tiene grado). Todo polinomio se expresa como suma de monomios.

La multiplicación por un escalar $\lambda \in K$, asocia al polinomio $P=[a_{ij}]$ el polinomio $\lambda P=[\lambda a_{ij}]$.

El conjunto de los polinomios adopta así la estructura de un

espacio vectorial de infinitas dimensiones; el conjunto de monomios de coeficientes iguales a 1 constituye una base del mismo.

3. Producto.—Se define en primer lugar el producto de dos monomios: $[(a)_{\alpha\beta}]$ y $[(b)_{\alpha'\beta'}]$ en la forma:

$$[(a)_{\alpha\beta}]\times[(b)_{\alpha'\beta'}]=[(ab)_{\alpha+\alpha',\,\beta+\beta'}].$$

Se hace el producto de los coeficientes y las sumas de los respectivos índices. El monomio $(1)_{00}$ es el elemento neutro; la conmutatividad y la asociatividad son evidentes.

Para multiplicar dos polinomios se consideran como sumas de monomios y se hace el producto de cada término del primero por cada término del segundo. Son también válidas aquí las demostraciones de la conmutatividad, asociatividad y distributividad con relación a la suma hechas en el caso de una variable.

Estas definiciones permiten utilizar la escritura habitual. Llamando x e y a los monomios

$$x=[0, 1, 0, 0, \ldots], \qquad y=[0, 0, 1, 0, 0, \ldots]$$

se demuestra que la base anterior es idéntica al conjunto de las potencias de los polinomios x e y:

$$x^2=[0, 0, 0, 1, 0, 0, \ldots], \qquad xy=[0, 0, 0, 0, 1, 0, 0, \ldots],$$
$$y^2=[0, 0, 0, 0, 0, 1, 0, 0, \ldots], \ \ldots \text{ etc.}$$

con

$$x^0=y^0=[1, 0, 0, 0, \ldots]$$

identificado con 1 en razón de un isomorfismo evidente. Un polinomio P de grado n se escribe entonces

$$P=a_{00}+a_{10}x+a_{01}y+a_{20}x^2+a_{11}xy+a_{02}y^2+\ldots+a_{0n}y^n.$$

Encontramos de nuevo la notación clásica y la técnica corriente de cálculo con sumas y productos.

El conjunto de los polinomios constituye un anillo de integridad, si suponemos que los coeficientes se eligen en un cuerpo: es el anillo $K[x, y]$ de los polinomios de dos variables sobre K.

De manera general, $K[x_1, x_2, \ldots, x_n]$ es el anillo de los polinomios de n variables sobre K.

OBSERVACIONES. 1.ª Es cómodo, muchas veces, considerar el anillo $K[x, y]$ como anillo de los polinomios en y cuyos coeficientes perte-

necen a K[x], es decir, K[x] [y]. En efecto, en virtud de las reglas de cálculo sobre K[x, y], se puede escribir

$$P = P_0 + P_1 y + P_2 y^2 + \ldots + P_n y^n,$$

donde

$$P_0 = a_{00} + a_{10}x + a_{20}x^2 + \ldots + a_{n0}x^n,$$
$$P_1 = a_{01} + a_{11}x + \ldots + a_{n1}x^n, \text{ etc.}$$

Se dice que P se ha ordenado con relación a y.

2.ª Esta manera de escribir un polinomio permite en ciertos casos la extensión de la división euclidiana (véase pág. 224). Así, p. ej., un polinomio cualquiera $P(x, y)$ puede dividirse por $y - \pi(x)$ puesto que el coeficiente de y es un elemento unitario.

En particular, la división se hace sin resto si, y solo si,

$$P[x, \pi(x)] = 0.$$

Igualmente, $P(x, y, z)$ es divisible por $z - \pi(x, y)$ cuando

$$P[x, y, \pi(x, y)] = 0.$$

3.ª Se puede asociar al anillo $K[x, y]$ el cuerpo de sus fracciones racionales, ya que se trata de un dominio de integridad.

4. Orden lexicográfico.—Sea un polinomio de tres variables x, y, z. El conjunto de los términos de grado determinado p se escribe, según los convenios anteriores,

$$P_p = \alpha x^p + \beta x^{p-1} y + \gamma x^{p-1} z + \delta x^{p-2} y^2 + \epsilon x^{p-2} yz + \eta x^{p-2} z^2 + \ldots$$

Decimos que los términos están dispuestos en *orden lexicográfico;* las diferentes palabras $xx \ldots x$ (p letras), $xx \ldots y$, $xx \ldots z$, ... se colocan como en un diccionario compuesto de palabras de p letras formadas con x, y, z.

Si $\alpha \neq 0$, decimos que αx^p es el primer término en el orden lexicográfico de P_p (en abreviatura p.t.o.l.).

Se demuestra fácilmente que haciendo el producto de dos polinomios homogéneos de las mismas variables, el p.t.o.l. es el producto de los dos primeros términos de los polinomios considerados.

EJERCICIOS:

1. Sobre C[x, y, z] efectúese el producto

$$(x + y + z)(x + jy + j^2 z)(x + j^2 y + jz).$$

2. Si a y b son dos elementos distintos de un cuerpo, redúzcanse los polinomios

$$P=\frac{1}{a^2(a^2-b^2)}(x^2-a^2)(y^2-b^2)+\frac{1}{b^2(b^2-a^2)}(x^2-b^2)(y^2-b^2)+\frac{x^2y^2}{a^2b^2}$$

$$Q=\frac{1}{a^2(a^2-b^2)}(x^2-a^2)(y^2-a^2)(z^2-a^2)+$$

$$+\frac{1}{b^2(b^2-a^2)}(x^2-b^2)(y^2-b^2)(z^2-b^2)+\frac{1}{a^2b^2}x^2y^2z^2.$$

3. Demuéstrese que si $P(x, y, z)$ es divisible separadamente por $x-y$, $y-z$, $z-x$, también lo es por su producto. Aplíquese al polinomio

$$P=xy^n+yz^n+zx^n-x^ny-y^nz-z^nx.$$

Precísese el cociente para $n=2$ y $n=3$.

4. Sobre el anillo $Q[x, y, z]$, descompónganse en factores de primer grado los polinomios:

$$P=(x-y)^3+(y-z)^3+(z-x)^3,$$
$$P_1=x^3(y-z)+y^3(z-x)+z^3(x-y).$$

5. Exprésese el polinomio

$$P=x^3(z-y^2)+y^3(x-z^2)+z^3(y-x^2)+xyz(xyz-1)$$

en forma de producto.

6. Simplifíquese la fracción racional

$$\frac{x^3(y^2-z^2)+y^3(z^2-x^2)+z^3(x^2-y^2)}{x^2(y-z)+y^2(z-x)+z^2(x-y)}$$

7. Efectúese la siguiente suma de fracciones:

$$\frac{x^3}{(x-y)(x-z)}+\frac{y^3}{(y-z)(y-x)}+\frac{z^3}{(z-x)(z-y)}$$

POLINOMIOS SIMETRICOS

5. Automorfismos definidos por sustituciones sobre las variables. Consideremos un anillo $K[X, Y]$. A todo polinomio $P(X, Y)$ hacemos corresponder el polinomio P' obtenido escribiendo Y en lugar de X y X en lugar de Y. Dicho de otra manera, $P'(X, Y)=P(Y, X)$. Evidentemente, $(P')'=P$. La correspondencia $P \Longleftrightarrow P'$ es un automorfismo del anillo $K[X, Y]$, ya que $P \Longleftrightarrow P'$ y $Q \Longleftrightarrow Q'$ implican de manera

manifiesta

$$(P+Q)' \Longleftrightarrow P'+Q',$$
$$(PQ)' \Longleftrightarrow P'Q'.$$

En general, en $K[X_1, X_2, \ldots, X_n]$ consideremos una sustitución s hecha sobre las letras $X_1, X_2, \ldots, X_n$ y sea $sX_i = X_i'$. Recíprocamente, $X_i = s^{-1}X_i'$. La correspondencia

$$P(X_1, X_2, \ldots, X_n) \Longleftrightarrow P' = P(X_1', X_2', \ldots, X_n')$$

es un automorfismo, porque es compatible con la igualdad, la suma y el producto en el anillo $K[X_1, X_2, \ldots, X_n]$.

Definición.— Los polinomios *invariantes* respecto al conjunto de los automorfismos anteriormente definidos se llaman *simétricos*.

Ya que una sustitución cualquiera es un producto de trasposiciones, basta comprobar su invariancia en todas las trasposiciones posibles para asegurarse de que un polinomio es simétrico. Ejemplos:

con n variables $P_1 = x_1^1 + x_2^1 + \ldots + x_n^1$,
con 4 variables $P_2 = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$.

Sea el monomio $x_1^\alpha x_2^\beta \ldots x_n^\lambda$ en el que los enteros $\alpha, \beta, \ldots, \lambda$ pueden ser nulos. Efectuando en él todas las sustituciones sobre las variables $x_1, x_2, \ldots, x_n$ se obtienen varios monomios *distintos* cuya suma es evidentemente un polinomio simétrico denotado por $\Sigma x_1^\alpha x_2^\beta \ldots x_n^\lambda$. Así, p. ej., el polinomio P_1 se escribe $\Sigma x_1^1 x_2^0 \ldots x_n^0$, siendo el número de términos distintos

$$\frac{n!}{(n-1)!} = n,$$

ya que las sustituciones sobre las $n-1$ últimas letras proporcionan el mismo monomio. Análogamente, $P_2 = \Sigma x_1x_2x_3^0x_4^0$ contiene

$$\frac{4!}{2! \times 2!} = 6 \text{ términos.}$$

Funciones simétricas fundamentales.—Sea el polinomio

$$P = (t-x_1)(t-x_2)\ldots(t-x_n),$$

en el que t designa una nueva variable. Desarrollando, resulta:

$$P = t^n - (\Sigma x_1)t^{n-1} + (\Sigma x_1x_2)t^{n-2} + \ldots + (-1)^p(\Sigma x_1x_2\ldots x_p)t^{n-p} + \ldots + (-1)^n x_1x_2\ldots x_n;$$

el coeficiente de t^{n-p} se obtiene tomando de todas las maneras posibles el término en t en $n-p$ factores y el término en x_i en los p factores restantes. Como una sustitución cualquiera sobre $x_1, x_2, \ldots, x_n$ no modifica P, los coeficientes de t en el desarrollo son ciertas funciones simétricas llamadas fundamentales:

$$A_1 = \Sigma x_1, \qquad A_2 = \Sigma x_1 x_2, \qquad \ldots, \qquad A_n = x_1 x_2 \ldots x_n.$$

Para cada una de ellas hemos escrito su p.t.o.l. Su importancia se desprende del siguiente teorema.

6. Teorema.—*Todo polinomio simétrico* P *de n variables* $x_1, x_2, \ldots, x_n$ *puede expresarse de manera única como polinomio en* $A_1, A_2, \ldots, A_n$.

Dado que cualquier sustitución sobre las variables no modifica el grado de un monomio, el conjunto de los términos de P de grado determinado, debe constituir un polinomio simétrico. Supondremos, pues, P homogéneo.

La demostración se articula así:

LEMA 1.—Si $\alpha x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n}$ es el primer término en el orden lexicográfico de P, se tiene que $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n$.

En efecto, si, p. ej., $a_2 > a_1$, hagamos la sustitución que permuta x_1 y x_2; el término considerado se transforma en otro término del polinomio simétrico P, a saber

$$\alpha x_1^{a_2} x_2^{a_1} \ldots x_n^{a_n}.$$

Pero en el orden lexicográfico este es anterior al p.t.o.l., lo que conduce a una contradicción.

LEMA 2.—Los monomios del tipo anterior y de un grado determinado forman un conjunto finito. En particular, en dicho conjunto existe un número bien determinado de elementos que son lexicográficamente posteriores a uno cualquiera de ellos.

LEMA 3.—Con la ayuda de los polinomios $A_1, A_2, \ldots, A_n$ es posible formar un polinomio simétrico que tenga por p.t.o.l. uno cualquiera de los monomios precedentes.

Se aplica la observación sobre el p.t.o.l. de un producto. El polinomio

$$(A_1)^{a_1 - a_2} (A_2)^{a_2 - a_3} \ldots (A_{n-1})^{a_{n-1} - a_n} (A_n)^{a_n}$$

tiene por primer término lexicográfico $x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n}$ (se supone $a_1 \geqslant a_2 \geqslant a_3 \ldots$).

Demostración de la posibilidad.—Sea $\mu_1 = \alpha_1 x_1^{a_1} x_2^{a_2}, \ldots, x_n^{a_n}$ el p.t.o.l. de P, y se forma el polinomio simétrico $\alpha_1 A_1^{a_1-a_2} A_2^{a_2-a_3} \ldots A_n^{a_n} = u_1$. La diferencia $P - u_1$ es un polinomio simétrico en el que ha desaparecido μ_1 (lema 3). Su p.t.o.l. μ_2 (importa poco que proceda de P o de u_1) sigue, pues, a μ_1 en el orden lexicográfico y es un término con exponentes no crecientes. Se comienza de nuevo y se obtiene la sucesión

$$P - u_1, \quad P - u_1 - u_2, \quad P - u_1 - u_2 - u_3, \ \ldots$$

tal que el p.t.o.l. de cada término resulta posterior al p.t.o.l. del término precedente de la sucesión. En virtud del lema 2, después de un número finito de operaciones se deberá necesariamente obtener el polinomio nulo. Tenemos así expresado P por medio de los polinomios A:

$$P(x_1, x_2, \ldots, x_n) = \varphi(A_1, A_2, \ldots, A_n).$$

Unicidad.—Supongamos que el polinomio P, simétrico con relación a $x_1, x_2, \ldots, x_n$, se expresa de dos maneras *diferentes* por medio de las funciones fundamentales $A_1, A_2, \ldots, A_n$:

$$P(x_1, x_2, \ldots, x_n) = \varphi(A_1, A_2, \ldots, A_n) = \psi(A_1, A_2, \ldots, A_n).$$

Entonces, la diferencia

$$\Delta(A_1, A_2, \ldots, A_n) = \varphi(A_1, A_2, \ldots, A_n) - \psi(A_1, A_2, \ldots, A_n)$$

es un polinomio cuyos coeficientes no son todos nulos; pero si reemplazamos las A_i por sus expresiones en función de las x_i, se obtiene un polinomio,

$$\Delta(x_1 + x_2 + \ldots + x_n, \ x_1x_2 + x_1x_3 + \ldots, \ \ldots, \ x_1x_2 \ldots x_n) = D(x_1, x_2, \ldots, x_n),$$

que se reduce al polinomio nulo de $K[x_1, x_2, \ldots, x_n]$.

Demostremos que esto es imposible, y que si el polinomio Δ no es nulo, tampoco lo es el polinomio D.

Sea $\alpha A_1^{a_1} A_2^{a_2} \ldots A_n^{a_n}$ un término de Δ. Expresando los A_i en función de los x_i se obtiene un polinomio cuyo p.t.o.l. es

$$\alpha x_1^{a_1} (x_1x_2)^{a_2} \ldots (x_1x_2 \ldots x_n)^{a_n} = \alpha x_1^{b_1} x_2^{b_2} \ldots x_n^{b_n},$$

con

$$\begin{cases} b_1 = a_1 + a_2 + a_3 + \ldots + a_n \\ b_2 = \quad a_2 + a_3 + \ldots + a_n \\ \cdot \\ \cdot \\ b_n = \quad a_n, \end{cases}$$

donde $a_i = b_i - b_{i+1}$ para $i = 1, 2, \ldots, n-1$ y $a_n = b_n$; de esta forma, términos distintos de Δ dan, después de reemplazadas las A_i, polinomios en x_i cuyos p.t.o.l. son todos diferentes. En particular, una vez transformados todos los términos de Δ, el primero en el orden lexicográfico de estos p.t.o.l. es el p.t.o.l. de D y no tiene término semejante, lo que demuestra que D no se reduce a O.

7. **Ejemplos.**—El método anterior puede aplicarse en la práctica; pero, en los casos más sencillos, es más rápido razonar directamente, con lo que se vuelve a encontrar el polinomio considerado en una igualdad que contiene $A_1, A_2, \ldots, A_n$.

1.° Calcúlese Σx_1^2.

El polinomio $A_1^2 = (x_1 + x_2 + \ldots + x_n)^2$ introduce el polinomio propuesto. Tenemos:

$$A_1^2 = \Sigma x_1^2 + 2A_2, \quad \text{de donde} \quad \Sigma x_1^2 = A_1^2 - 2A_2.$$

2.° Calcúlese Σx_1^3.

Se parte de la igualdad

$$(\Sigma x_1^2) \times (\Sigma x_1) = \Sigma x_1^3 + \Sigma x_1 x_2^2,$$

observando que el último polinomio de ella, $\Sigma x_1 x_2^2$, se calcula fácilmente:

$$\Sigma x_1 x_2^2 = (\Sigma x_1 x_2) \times (\Sigma x_1) - 3\Sigma x_1 x_2 x_3.$$

Se deduce:

$$\Sigma x_1^3 = A_1^3 - A_1 A_2 + 3A_3.$$

Para el cálculo de Σx_1^m por recurrencia, véase el ejercicio 13-5 (pág. 285).

3.° Calcúlese $\Sigma x_1^2 x_2 x_3$.

El método del texto nos conduce a formar $A_1 A_3$. Se obtiene:

$$A_1 A_3 = (\Sigma x_1)(\Sigma x_1 x_2 x_3) = \Sigma x_1^2 x_2 x_3 + 4\Sigma x_1 x_2 x_3 x_4,$$

de donde,

$$\Sigma x_1^2 x_2 x_3 = A_1 A_3 - 4A_4.$$

4.º Sea un polinomio homogéneo y simétrico de grado d. Ya que las funciones A_i son también homogéneas y de grado i, la expresión de P será de la forma

$$P = \Sigma A_1^{b_1} A_2^{b_2} \dots A_n^{b_n}$$

cumpliéndose además la relación $b_1 + 2b_2 + \dots + nb_n = d$.

De esta forma se puede simplificar el proceso, precisando los valores posibles de $b_1, b_2, \dots, b_n$.

Para cuatro variables de suma nula x_1, x_2, x_3, x_4, calcúlese la función simétrica [1]

$$D = \prod_{i \neq k} (x_i - x_k).$$

Es homogénea y de grado 12; de donde, $12 = b_1 + 2b_2 + 3b_3 + 4b_4$ y, como $A_1 = 0$,

$$2b_2 + 3b_3 + 4b_4 = 12.$$

El p.t.o.l. de D es $x_1^6 x_2^4 x_3^2$, lo que demuestra que A_2^6 no figura en D, siendo su p.t.o.l. $x_1^6 x_2^6$, que es anterior al de D. De ello resulta una relación de la forma

$$D = \xi_1 A_2^4 A_4 + \xi_2 A_2^3 A_3^2 + \xi_3 A_2^2 A_4^2 + \xi_4 A_2 A_3^2 A_4 + \xi_5 A_3^4 + \xi_6 A_4^3;$$

para calcular los coeficientes ξ_i se reemplazan los x_i por valores sencillos que verifiquen $A_1 = 0$. Hagamos el cálculo en $K = Q(i)$.

a) $x_1 = 1, \quad x_2 = -1, \quad x_3 = i, \quad x_4 = -i$ dan $P = x^4 - 1$,

$A_1 = A_2 = A_3 = 0, \quad A_4 = -1, \quad D = -256,$ de donde $\xi_6 = 256.$

b) $x_1 = x_2 = 1, \quad x_3 = x_4 = -1$ y $x_1 = 1, \quad x_2 = -1, \quad x_3 = 2, \quad x_4 = -2$

dan, después de algunos cálculos, que efectuará el lector, y teniendo en cuenta que $\xi_6 = 256$,

$$4\xi_1 + \xi_3 + 64 = 0, \qquad 25\xi_1 + 4\xi_3 + 112 = 0,$$

de donde $\xi_1 = 16$ y $\xi_3 = -128$.

[1] La letra Π indica el producto de todos los términos análogos a $x_i - x_k$, cuando i y k toman todos los valores de 1 a 4, con $i \neq k$.

c) Para tener $A_4=0$, tomemos $x_1=x_2=1$, $x_3=-2$, $x_4=0$, después $x_1=1$, $x_2=2$, $x_3=-3$, $x_4=0$. Se obtiene:

$$27\xi_2-4\xi_5=0, \qquad 343\xi_2-36\xi_5+400=0,$$

de donde $\xi_2=-4$, $\xi_5=-25$.

d) $x_1=x_2=x_3$ y $x_4=-3$ dan, finalmente, $\xi_4=+144$.

EJERCICIOS:

1. Si a, b, c designan las tres raíces de la ecuación $x^3+px+q=0$, exprésense en función de p y q las funciones simétricas

$$\Sigma a^2b^2, \qquad (a-b)^2(b-c)^2(c-a)^2.$$

2. Exprésese el polinomio simétrico $\Sigma(x_1+x_2)(x_1+x_3)(x_1+x_4)$ de las cuatro variables x_1, x_2, x_3, x_4 mediante las funciones simétricas fundamentales.

3. Calcúlese la función $\Sigma\dfrac{1}{x_i}$ para n variables. (Efectuando operaciones se reduce al cociente de dos polinomios simétricos.)

4. Calcúlense igualmente

$$\Sigma\frac{x_1}{x_2}, \qquad \Sigma\frac{x_1^2}{x_2}, \qquad \Sigma\frac{x_1}{x_2^2}$$

utilizando el resultado del ejercicio anterior.

5. Calcúlese por el método dado en la demostración general el polinomio $\Sigma x_1^4x_2^2x_3^2x_4^1$.

6. Exprésense con ayuda de A_1, A_2, A_3, A_4 los polinomios simétricos siguientes:

$$(x_1+x_2)(x_1+x_3)(x_1+x_4)(x_2+x_3)(x_2+x_4)(x_3+x_4);$$
$$(x_1x_2+x_3x_4)(x_1x_3+x_2x_4)(x_1x_4+x_2x_4);$$
$$(x_1+x_2-x_3-x_4)(x_1-x_2+x_3-x_4)(x_1-x_2-x_3+x_4).$$

Resp.: $A_1A_2A_3-A_1^2A_4-A_3^2$; $A_1^2A_4+A_3^2-4A_2A_4$; $A_1^4-4A_1A_2+8A_3$.

7. Sea $x^3+ax^2+bx+c=0$. Hállese la relación que debe existir entre a, b, c para que las raíces sean los senos de los ángulos de un triángulo.

Resp.: $a(4ab-a^3-8c)-4c^2=0$.

FACTORIZACION UNICA EN K[x, y]

8. Anillo de factorización única. **Definición.**—Un dominio de integridad lo es de factorización única si se cumplen las dos condiciones que siguen:

a) todo elemento no unitario puede descomponerse en factores primos;

b) esta descomposición es única si consideramos como idénticas dos descomposiciones que difieran en el orden de los factores o por la presencia de factores unitarios. El anillo Z de los enteros relativos y el anillo K[x] de los polinomios de una variable son de factorización única. Estos resultados se han obtenido a partir de la igualdad de Bezout (véase pág. 233), que da la condición necesaria y suficiente para que dos elementos A y B sean primos entre sí:

$$SA + TB = 1.$$

Esta igualdad no es cierta sobre el anillo K[x, y] de los polinomios de dos variables. Así, p. ej., si K es el cuerpo de los racionales, los polinomios x e $y^2 - x$ no tienen divisores comunes (a excepción de los elementos unitarios), a pesar de que es imposible encontrar dos polinomios S(x, y) y T(x, y) tales que

$$S \times x + T \times (y^2 - x) = 1,$$

porque el primer miembro carece de término constante, cualquiera que sea la elección de S y de T.

Sin embargo, vamos a demostrar que el anillo K[x, y] es de factorización única. Estudiemos en primer lugar un caso más conocido del lector.

9. Caso del anillo Z[x].—Consideremos el anillo Z[x] de los polinomios con coeficientes enteros. La igualdad de Bezout no es válida; sin embargo, demostraremos que el anillo es de factorización única. Un polinomio cualquiera

$$P(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \ (a_0, a_1, \ldots, a_n, \text{ enteros})$$

puede escribirse en la forma

$$P(x) = \lambda(a'_0x^n + a'_1x^{n-1} + \ldots + a'_n),$$

llamando λ al M. C. D. de los coeficientes: λ está definido, salvo un factor unitario, es decir, salvo el signo, sobre Z. El polinomio entre paréntesis, cuyos coeficientes son primos entre sí, se llama *primitivo*. Se ve que los polinomios irreducibles de Z[x] son, por una parte, los números primos, y, por otra, los polinomios primitivos no descomponibles; p. ej., $x + 1$, $x^2 - 2$, etc.

La propiedad que se desea demostrar resulta del teorema que sigue:

TEOREMA.—*Un polinomio de coeficientes enteros, descomponible en un producto de polinomios con coeficientes racionales, puede igualmente factorizarse en polinomios de los mismos grados con coeficientes enteros.*

Demostremos dos lemas previos.

LEMA 1.—*El producto de dos polinomios primitivos es primitivo.*

Los dos polinomios

$$P(x)=\sum_0^n a_i x^i \qquad \text{y} \qquad Q(x)=\sum_0^p b_j x^j$$

tienen por producto

$$R(x)=\sum_0^{n+p} c_k x^k, \qquad \text{con} \qquad c_k=\sum_{i+j=k} a_i b_j.$$

Supongamos $R(x)$ no primitivo: un mismo número primo p divide a todos los c, pero todos los a y todos los b no pueden ser divisibles por p. Sean, pues, en el orden creciente de los índices, a_λ y b_μ los primeros coeficientes no divisibles. Se tiene:

$$a_\lambda b_\mu = c_{\lambda+\mu}(a_0 b_{\lambda+\mu}+a_1 b_{\lambda+\mu-1}+\ldots+a_{\lambda-1}b_{\mu+1}+a_{\lambda+1}b_{\mu-1}+\ldots+a_{\lambda+\mu}b_0).$$

Esta igualdad implicaría $p|a_\lambda b_\mu$ y, por consiguiente, $p|a_\lambda$ o $p|b_\mu$, lo que es contradictorio.

LEMA 2.—*Todo polinomio no nulo $f(x)$ de coeficientes racionales ($f(x)\in Q[x]$) puede escribirse $f(x)=\lambda f^*(x)$, donde $f^*(x)$ es primitivo y λ es un número racional. La descomposición obtenida es única salvo factores unitarios.*

El polinomio de coeficientes racionales

$$f(x)=\frac{a_n}{b_n}x^n+\frac{a_{n-1}}{b_{n-1}}x^{n-1}+\ldots+\frac{a_0}{b_0}$$

puede escribirse:

$$f=\frac{1}{b_n b_{n-1}\ldots b_0}f_1(x),$$

donde $f_1(x)$ tiene coeficientes enteros. Si p es su M. C. D., tenemos

$$f=\frac{p}{b_n b_{n-1}\dots b_0}f^*(x)=\lambda f^*(x),$$

siendo el polinomio f^* primitivo.

Si hubiera una segunda descomposición, $f=\lambda' g^*(x)$, se deduciría que $\lambda f^*(x)=\lambda' g^*(x)$ y, multiplicando ambos miembros por un mismo número racional que depende de λ y de λ', se tendría

$$\mu f^*=\mu' g^*,$$

donde μ y μ' son dos enteros primos entre sí. El entero μ', que divide a todos los coeficientes del polinomio μf^*, debe dividir a su M. C. D. μ; puesto que μ y μ' son primos entre sí, esto sólo es posible si μ y μ' son elementos asociados (iguales salvo el signo sobre Z); f^* y g^* lo son también.

Demostremos ahora el teorema enunciado al principio del apartado; supongamos f de Z[x] descompuesto según

$$f=g\cdot h \quad \text{donde} \quad g \text{ y } h \text{ pertenecen a } Q[x].$$

Tenemos $g=\lambda g^*$ y $h=\lambda_1 h^*$, de donde $f=\lambda\lambda_1 g^* h^*$; pero también, en Z[$x$], $f=\nu f^*$, siendo ν entero. En virtud del lema I, el polinomio g^*h^* es primitivo, y es asociado a f^* según el lema 2; por consiguiente, ν es asociado de $\lambda\lambda_1$ y, por tanto, $f=\nu g^* h^*$, que constituye una descomposición de f con coeficientes enteros. Los polinomios g^* y h^* tienen además los mismos grados que f y g.

De lo anterior se deduce el teorema que sigue:

TEOREMA.—*El anillo* Z[x] *de los polinomios de coeficientes enteros es de factorización única.*

Todo polinomio f de Z[x], considerado como perteneciente a Q[x], admite una descomposición única [1], $f=ghk$. Hemos visto que se puede escribir $f=\nu g^* h^* k^*$, siendo ν un entero expresable como producto de factores primos. Los polinomios primitivos g^*, h^*, k^*, irreducibles en Q[x], lo son también en Z[x]: son factores primos. La descomposición obtenida es única (salvo los factores unitarios), en virtud de la correspondencia entre los elementos irreducibles en Z[x] y en Q[x].

10. Generalización.—El razonamiento que acabamos de hacer en el anillo Z[x] es en realidad mucho más general y se aplica a todo do-

[1] Que puede comprender un solo factor.

minio de integridad D de factorización única: también el anillo $D[x]$ de los polinomios con una variable sobre D es de factorización única.

Esta circunstancia pone además de manifiesto la riqueza de la idea de estructura: el mismo razonamiento abstracto se aplica a casos concretos muy diversos. No repetiremos el razonamiento general, pero vamos a dar algunas precisiones para el caso que nos ocupa: el anillo $K[x, y]$ con dos variables sobre un cuerpo K. Hemos visto que lo podemos considerar como el anillo $K[x][y]$, con una variable y, cuyos coeficientes pertenecen al anillo $K[x]$. Siendo este último un dominio de integridad, de factorización única, es aplicable el método precedente. Un polinomio de $K[x][y]$ es primitivo cuando todos sus coeficientes son primos entre sí en $K[x]$; así, p. ej.,

$$(3x^2+1)y+(x+2).$$

Junto a los polinomios $K[x][y]$, introducimos los polinomios en y, $K(x)[y]$, cuyos coeficientes son fracciones racionales en x; el lema 2 se aplica idénticamente, observando que los factores unitarios son ahora los elementos de K. De ello se deduce que todo polinomio con dos variables, descomponible en $K(x)[y]$, lo es ya en $K[x][y]$ y se termina como anteriormente.

El mismo método permite demostrar que el anillo $K[x, y, z]$ es también de factorización única.

EJERCICIOS SOBRE EL CAPITULO XIII

13-1. Demuéstrese que si

$$(xy+yz+zx)^3=xyz(x+y+z)^3,$$

los tres números complejos x, y, z están en progresión geométrica. Considérense x, y, z como raíces de una ecuación de tercer grado.

13-2. Sobre el cuerpo de los reales resuélvase el sistema

$$\begin{cases} y^2+yz+z^2=a^2, \\ z^2+zx+x^2=b^2, \\ x^2+xy+y^2=c^2. \end{cases}$$

Tómese como incógnita auxiliar $m=x+y+z$. Exprésense x, y, z con ayuda de m (y de a, b, c). Dedúzcase de ello una ecuación bicuadrada a la que satisfaga m.

13-3. Sobre el cuerpo de los racionales, fórmese una ecuación de tercer grado cuyas raíces x, y, z verifiquen el sistema

$$\begin{cases} xyz=6, \\ x^2+y^2+z^2=14, \\ x+y+z=6. \end{cases}$$

Dedúzcase de ello x, y, z.

13-4. Redúzcase la resolución del sistema

$$\begin{cases} x^3+y^3+z^3=a, \\ x^4+y^4+z^4=b, \\ xy+yz+zx=c, \end{cases}$$

a la de una ecuación de tercer grado.

13-5. *Fórmulas de Newton.* Permiten el cálculo de las sumas de potencias semejantes de n variables: $S_k=\Sigma x_1^k$.

1.° Transfórmese el segundo miembro y compárese al polinomio $S_{k-1}\cdot S_1$ de igual p.t.o.l. Se halla

$$S_k=S_{k-1}A_1-\Sigma x_1^{k-1}x_2.$$

2.° Se repite la aplicación de este método al último término de la fórmula precedente. Después de p operaciones se llega, suponiendo p más pequeño que k y que n, a

$$S_k=S_{k-1}A_1-S_{k-2}A_2+\ldots+(-1)^{p-1}S_{k-p}A_p+(-1)^p\Sigma x_1^{k-p}x_2\ldots x_{k-p}$$

3.° Prosiguiendo, pueden presentarse dos casos:
$k\leqslant n$; se tiene:

$$S_k=S_{k-1}A_1-S_{k-2}A_2+\ldots+(-1)^{k-2}S_1A_{k-1}+(-1)^{k-1}kA_k,$$

$k>n$

$$S_k=S_{k-1}A_1-S_{k-2}A_2+\ldots+(-1)^{n-1}A_nS_{k-n}.$$

4.° Para n variables, calcúlense por las fórmulas anteriores S_4 y S_5 según la posición de n con relación a los enteros 4 y 5.

13-6. Si un polinomio primitivo $A(x)\in Z[x]$, de grado n, puede descomponerse en la forma $A(x)=B(x)\cdot C(x)$ (B y C con coeficientes enteros), existe la misma relación entre los valores de A, B y C para todo valor entero de la variable. Dedúzcase de ello que los polinomios de grado $l<n$ que dividen a A se pueden elegir entre un número finito de polinomios, que se construyen dando $l+1$ valores distintos [1] a x.

Indíquese un método (poco eficaz en la práctica) para reconocer si un polinomio de $Q[x]$ es o no descomponible.

13-7. *Criterio de Eisenstein.*—Sea $C(x)=c_0+c_1x+\ldots+c_nx^n$ un polinomio primitivo de coeficientes enteros, y p, un número primo que divide a c_0, c_1, ..., c_{n-1} sin dividir a c_n. Demuéstrese que si $C(x)$ es descomponible en un producto:

$$C(x)=(a_0+a_1x+\ldots+a_qx^q)(b_0+b_1x+\ldots+b_rx^r),$$

el número p divide a todos los coeficientes a_i y b_j, a excepción de a_q y b_r. Conclúyase de lo anterior que c_0 es divisible por p^2.

Enúnciese una condición suficiente para que $C(x)$ sea indescomponible sobre $Q(x)$.

---

[1] Un polinomio de grado l está determinado si se dan sus valores para $l+1$ valores distintos de la variable.

Ejemplo.—Sea $C(x)=x^{p-1}+x^{p-2}+\ldots+x+1$. Fórmese el polinomio $P(y)=C(y+1)$ observando que $C(x)=\dfrac{x^p-1}{x-1}$ y demuéstrese que es aplicable el criterio de Eisenstein.

EJERCICIOS DE RECAPITULACION (Partes II y III)

1. Compruébese la identidad

$$(a+b+c)^4-2(a^2+b^2+c^2)(a+b+c)^2-8abc(a+b+c)+ \\ +(a^2+b^2+c^2)^2-4(b^2c^2+c^2a^2+a^2b^2)=0$$

y dedúzcase un método para reducir $z^4+pz^2+qz+r=0$ a una ecuación de tercer grado.

2. Se consideran los polinomios A, B, C enteros en z, con coeficientes reales o complejos, primos dos a dos y que satisfacen la identidad $A^2+B^2=C^2$.

a) Demuéstrese que los polinomios $C+B$ y $C-B$ son cuadrados perfectos y dése, en consecuencia, la forma general de los polinomios A, B, C.

b) Determínense los polinomios A y B en el caso de $C=z^2+a^2$. Demuéstrese que, si a es real y no nulo y si A y B tienen sus coeficientes reales, son de la forma

$$A=(z^2-a^2)\operatorname{sen}\alpha+2az\cos\alpha,$$
$$B=(z^2-a^2)\cos(\alpha+k\pi)-2az\operatorname{sen}(\alpha+k\pi),$$

designando α una constante cualquiera y k un entero arbitrario.

(Ecoles des Ponts et Chaussées et des Télécommunications, 1950.)

3. *a)* Sea el polinomio

$$f(z)=z^3+\frac{1+i\sqrt{7}}{2}z^2-\frac{1-i\sqrt{7}}{2}z-1.$$

Demuéstrese que $f(z^2)$ es divisible por $f(z)$. ¿Qué relación existe entre el cociente $g(z)$ de la división y $f(z)$?

b) Sea a una raíz de la ecuación $f(z)=0$; pruébese que $b=a^2$ y $c=b^2$ son también raíces de esta ecuación. ¿Pueden ser iguales dos de los tres números a, b, c? Demuéstrese que $c^2=a$ y dedúzcase de ello las tres raíces de la ecuación.

c) Estúdiense los restos de las divisiones por $f(z)$ de los monomios z, z^2, z^4, ..., que se deducen los unos de los otros reemplazando z por z^2 en el monomio precedente. Si R_1, R_2, R_3 designan los tres primeros restos, pruébese que se obtiene una sucesión periódica, repitiéndose indefinidamente los tres precedentes, en este orden.

4. Dados los números a, h, b, b', c, c', con $h\neq 0$, hállese un polinomio $f(x)$ que satisfaga las cuatro condiciones

$$f(a)=b, \quad f(a+h)=c, \quad f'(a)=b', \quad f'(a+h)=c'. \qquad [1]$$

1.° Se supone que el polinomio $f(x)$ es de tercer grado; haciendo

$$f(x)\equiv\alpha x^3+\beta x^2+\gamma x+\delta,$$

fórmese el sistema de ecuaciones lineales a que deben satisfacer α, β, γ, δ y demuéstrese que α, β, γ, δ están determinados de manera única.

2.° Calcúlese α: dedúzcase de lo anterior que la condición

$$c-b=h\,\frac{b'+c'}{2} \qquad [2]$$

es necesaria y suficiente para que exista un polinomio de segundo grado que verifique [1].

Sobre la curva C, representativa de la función $y=f(x)$, sean A, M los puntos de abscisas a, $a+h$, y sea T la intersección de las tangentes en A y M a C; calcúlese la abscisa t de T en función de a, h, b, b', c, c'. Demuéstrese que la relación [2] expresa que t es media aritmética entre a y $a+h$.

¿Qué propiedad de la parábola se obtiene así?

3.° Se supone que el polinomio $f(x)$ es de grado cualquiera; demuéstrese que, para que verifique [1], es necesario y suficiente que se pueda escribir en la forma $f_3(x)+(x-a)^2(x-a-h)^2\,P(x)$, siendo $f_3(x)$ el polinomio de tercer grado determinado en el caso 1.°, y $P(x)$, un polinomio arbitrario.

5. *Anillos euclidianos.*—Sea E un dominio de integridad en el que se verifican las condiciones siguientes:

α) a todo elemento a corresponde un entero positivo o nulo $g(a)$;

β) $g(0)=0$, $g(a)\neq 0$ para $a\neq 0$;

γ) $g(ab)=g(a)g(b)$;

δ) dados dos elementos a y b cualesquiera, es siempre posible hallar otros dos q y r tales que

$$a=bq+r, \qquad g(r)<g(b).$$

E se llama un anillo euclidiano.

1.° Demuéstrese que E es un dominio de factorización única y que todo ideal está formado por los múltiplos de cierto elemento a tal que $g(a)$ sea mínimo.

2.° ¿Cuáles son los anillos euclidianos ya estudiados?

3.° *a)* Demuéstrese que $Z[\sqrt{2}]$ es un anillo euclidiano. Se definirá la función g a partir de la norma N del elemento

$$N(\alpha+\beta\sqrt{2})=(\alpha+\beta\sqrt{2})(\alpha-\beta\sqrt{2})=\alpha^2-2\beta^2.$$

b) Descompóngase en factores primos 5, 7, $2+\sqrt{2}$; para los elementos unitarios del anillo se referirá al problema 5-8 (pág. 97).

4.° El cuerpo $Q[x]/(x^2+x+1)$ es isomorfo a la extensión de Q por adjunción de un elemento j tal que $j^2+j+1=0$. Demuéstrese que el anillo $Z[j]$ es euclidiano. La norma de $a+bj$ es $(a+bj)(a+bj^2)=a^2-ab+b^2$.

5.° El cuerpo $Q[x]/(x^2+5)$ es isomorfo a la extensión de Q por adjunción de un elemento ω tal que $\omega^2+5=0$. Demuéstrese que el anillo $Z[\omega]$ no es euclidiano. Utilícense las igualdades $9=3\times 3=(2+\omega)(2-\omega)$.

6. *Enteros de Gauss.*—Se consideran los números complejos $a+bi$, donde a y b son *enteros* relativos.

1.° Pruébese que constituyen un anillo $Z[i]$.

2.° Se designa por $N(z)=a^2+b^2$ la norma de $z=a+bi$. Demuéstrese que $N(Z+Z_1)\leqslant N(Z)+N(Z_1)$.

3.° Dados dos enteros de Gauss, α y β, demuéstrese que se pueden hallar otros dos, χ y ρ, tales que

$$\alpha=\beta+\rho\chi, \qquad N(\rho)<N(\beta).$$

Se puede efectuar la división exacta de α por β en el cuerpo de los complejos y elegir el entero de Gauss cuyas componentes se aproximen más a las del cociente.

4.° Pruébese que el algoritmo de Euclides se aplica en el anillo de los enteros de Gauss.

Dedúzcase de lo anterior la existencia para α y β de un m. c. d., Δ, definido salvo un factor unitario. Dedúzcase igualmente la igualdad de Bezout.

5.° Demuéstrese que la descomposición en factores primos es única (salvo los factores unitarios). Dedúzcanse de ello la existencia de un m. c. m., μ, la relación $ab\sim\mu\Delta$, etc.

6.° Descompóngase en factores primos los enteros 2, 5, $5+i$, 11.

Determínese el m. c. d. de $19+13i$ y $23-9i$.

7.° Demuéstrese que un número primo ordinario p es primo en $Z[i]$, si, y solo si, la ecuación $x^2+y^2=p$ carece de soluciones enteras.

8.° Pruébese que todo ideal de $Z[i]$ está constituido por los múltiplos de un elemento de norma mínima.

PARTE IV

ALGEBRA LINEAL

CAPITULO XIV

LAS MATRICES

NOCIONES GENERALES Y NOTACIONES

1. Definición.—Se llama *matriz* de dimensiones m y n sobre un dominio $\mathfrak{D}$ un cuadro rectangular de m *filas* y n *columnas*, formado con elementos de $\mathfrak{D}$. El elemento situado en la intersección de la fila i con la columna j se designará por a_i^j o a_j^i, o a_{ij}. En el presente capítulo y salvo que se advierta explícitamente otra cosa, el subíndice designará la fila, y el superíndice, la columna. Los elementos de las matrices que consideraremos pertenecerán en general a un cuerpo conmutativo $\mathbb{K}$; es el caso más importante para las aplicaciones.

El cuadro puede ser, eventualmente, cuadrado—la matriz se denomina entonces cuadrada—y debe distinguirse de un determinante (véase pág. 376). Por ello se encierra entre corchetes alargados. (También se utilizan paréntesis grandes y dobles barras.) El ente matemático definido de esta manera se designa con la letra $\mathbf{A}$; escribiremos, pues,

$$\mathbf{A} = \begin{bmatrix} a_1^1 & a_1^2 & \dots & a_1^n \\ a_2^1 & a_2^2 & \dots & a_2^n \\ \dots & \dots & \dots & \dots \\ a_m^1 & a_m^2 & \dots & a_m^n \end{bmatrix} \quad \text{o, abreviadamente,} \quad [a_i^j].$$

Llamaremos *altura* de la matriz al número m de sus filas, y *longitud*, al número n de sus columnas.

Igualdad.—La igualdad de dos matrices se define por su identidad: $\mathbf{A}=\mathbf{B}$ si, y solo si, $a_i^j=b_i^j$ para todo $i=1, 2, \dots, m$ y todo $j=1, 2, \dots, n$. Implica, por tanto, la igualdad entre las dimensiones.

Cálculo matricial.—Las matrices de dimensiones finitas sobre un dominio forman un conjunto, y las matrices de $m \times n$ elementos, o matrices (m, n), un subconjunto del precedente. En este conjunto, o en alguna de sus partes, se definen leyes de composición tanto internas como externas, así como diversas relaciones y algoritmos; esto constituye el *cálculo matricial*. Las primeras aplicaciones que daremos de

él demostrarán la utilidad y fecundidad de las nociones introducidas aquí *a priori*.

Comencemos por dar algunos ejemplos de matrices.

1.° En un espacio vectorial de n dimensiones, un vector cualquiera puede representarse por una matriz después de haber elegido una base: los elementos de la matriz son las componentes del vector; se pueden escribir bien en fila, o bien en columna:

$$X=[x^1 \quad x^2 \quad \dots \quad x^n] \qquad \text{o} \qquad X=\begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}.$$

El lector observará, sin embargo, que un vector no es una matriz: un cambio de base cambia las componentes y, por tanto, la matriz, mientras el vector queda invariable.

2.° La teoría de las *sustituciones lineales* ha dado origen al cálculo matricial. Se efectúa una sustitución lineal sobre un conjunto ordenado de n elementos $(x_1, x_2, \dots, x_n)$ de un cuerpo K, cuando se les asocia m elementos $(y_1, y_2, \dots, y_m)$ definidos por las fórmulas siguientes:

$$\begin{cases} y_1=a_1^1x_1+a_1^2x_2+\dots+a_1^nx_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_m=a_m^1x_1+a_m^2x_2+\dots+a_m^nx_n, \end{cases} \quad (a_i^j \in \mathrm{K})$$

La sustitución queda caracterizada por una matriz (m, n)

$$A=\begin{bmatrix} a_1^1 & \dots & a_1^n \\ \vdots & & \vdots \\ a_m^1 & \dots & a_m^n \end{bmatrix}.$$

La teoría de las sustituciones está actualmente reemplazada por la de las aplicaciones lineales (véase Cap. XV). Sin embargo, la utilizamos en la sección 12 para justificar la regla de multiplicación de matrices.

3.° Las matrices de una fila y una columna contienen un único elemento; existe correspondencia biunívoca, $[a] \Longleftrightarrow a$, entre estas matrices y los elementos del dominio $\mathfrak{D}$. Si $\mathfrak{D}$ es un cuerpo, las operaciones entre matrices harán de esta correspondencia un isomorfismo, lo que permite la asimilación entre $[a]$ y a. De este modo, el cálculo matricial aparece como una amplia generalización del cálculo con los números.

4.° Consideremos dos conjuntos finitos $E=\{a_1, a_2, \ldots, a_m\}$ y $F=\{b_1, b_2, \ldots, b_n\}$ y una relación $\mathfrak{R}$ definida entre los a_i y los b_j. Hagamos $\alpha_{ij}=1$, si y solo si, a_i y b_j verifican $\mathfrak{R}$; $\alpha_{ij}=0$, en caso contrario. El lógico llama a $[\alpha_{ij}]$ *la matriz de verdad* de la relación $\mathfrak{R}$.

2. Subdivisión en cajas.—Los elementos de una matriz intervienen con mayor frecuencia en los cálculos como intermediarios de ciertos conjuntos que como seres individuales.

Si trazamos paralelas a los bordes del cuadro rectangular, la matriz se divide en cierto número de *submatrices* o *cajas*. Así, p. ej., la matriz A, que figura más abajo, se ha dividido en 6 cajas. Inversamente, se puede considerar A como la yuxtaposición de 6 cajas en un determinado orden, bien entendido que las escritas en una misma fila tienen necesariamente la misma altura, y las de una misma columna poseen la misma longitud. Se escribirá, pues, abreviadamente, $A=[B_j^i]$, lo que generaliza $A=[a_j^i]$:

$$A=\left[\begin{array}{ccc:c:ccc} a & b & c & d & e & f & g \\ a' & b' & c' & d' & e' & f' & g' \\ a'' & b'' & c'' & d'' & e'' & f'' & g'' \\ \hdashline \alpha & \beta & \gamma & \delta & \epsilon & \zeta & \eta \\ \alpha' & \beta' & \gamma' & \delta' & \epsilon' & \zeta' & \eta' \end{array}\right];$$

$$A=\begin{bmatrix} B_1^1 & B_1^2 & B_1^3 \\ B_2^1 & B_2^2 & B_2^3 \end{bmatrix}$$

Esta manera de escribir A muestra que subdividir en submatrices equivale a considerar A como una matriz sobre un dominio $\mathfrak{D}$ cuyos elementos son, ellos mismos, matrices.

La última descomposición se obtiene considerando la matriz como formada por la yuxtaposición de mn matrices de un elemento. En virtud de la asimilación ya señalada, $[a]=a$, volvemos de nuevo a la definición inicial.

Matriz fila y matriz columna.—Una matriz que contenga n elementos dispuestos en una fila, es decir, una matriz $(1, n)$, se llama *matriz fila*. Igualmente se define la *matriz columna*.

Después de lo dicho, una matriz se puede considerar como la yuxtaposición en hileras ordenadas verticalmente de matrices columna, o como la yuxtaposición apilada ordenadamente de matrices fila. Las cajas de este género son de empleo constante.

Dada una matriz (m, n), A, se simboliza por A_i la fila i, y con A^j la columna j. Con esta notación podemos escribir

$$A=\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} \qquad \text{y} \qquad A=[A^1\ A^2 \dots A^n].$$

Vectores fila y vectores columna.—Sea A una matriz (m, n). Toda fila está formada por n elementos ordenados de K. Por tanto, le corresponde un vector bien determinado del espacio K^n referido a la base canónica $e_1=(1, 0, 0, \dots, 0)$, $e_2, \dots, e_n$. Denotemos por $A_1, A_2, \dots, A_m$ los m vectores obtenidos de esta forma, llamados vectores fila, que engendran el subespacio fila de la matriz. Se define de manera análoga los vectores columna, $A^1, A^2, \dots, A^n$, y el subespacio correspondiente.

El lector encontrará, quizá, superabundantes estas definiciones de matrices fila y vectores fila. Existe, sin embargo, una distinción: las sucesiones ordenadas de n elementos de un cuerpo K constituyen el espacio K^n; pero, el orden puede obtenerse de muchas maneras y el cálculo matricial considera dos órdenes particulares: ordenar los elementos por filas o por columnas. Este cálculo considera, pues, como distintos el conjunto de matrices columna de altura n y el conjunto de las filas de la misma longitud, a pesar de que estos dos conjuntos sean isomorfos con K^n. En todas aquellas cuestiones donde solo intervenga la estructura de espacio vectorial de las filas o de las columnas, podemos limitarnos a considerar los vectores A_i o A^j, que les están asociados.

3. Trasposición.—En el conjunto de las matrices se define una operación unaria llamada *trasposición*. Hallar la traspuesta de una matriz (m, n), A, es formar una matriz (n, m), ${}^tA$, que tiene como primera fila la primera columna de A, por segunda fila la segunda columna de A, etc. (se respeta el orden). Ejemplo:

$$A=\begin{bmatrix} 3 & 2 & 1 \\ 2 & 0 & 8 \end{bmatrix} \qquad {}^tA=\begin{bmatrix} 3 & 2 \\ 2 & 0 \\ 1 & 8 \end{bmatrix}.$$

OBSERVACIONES.—1.ª Se puede decir también que ${}^tA$ admite por sucesivas columnas las filas sucesivas de A.

2.ª Cualquiera que sea A, se tiene siempre

$${}^t({}^tA)=A.$$

3.ª De manera general, $A=[a_j^i]$ se convierte en ${}^tA=[a_i'^j]$, con $a_i'^j=a_j^i$.

4.ª Una matriz que coincide con su traspuesta es, necesariamente, cuadrada; se denomina *simétrica* y se verifica que $a_i^j=a_j^i$.

ESPACIO VECTORIAL DE LAS MATRICES DE DIMENSIONES DADAS

Las definiciones que damos a continuación tienen por objeto dotar al conjunto de las matrices (m, n) sobre el cuerpo K, de una estructura de *espacio vectorial*.

4. Adición.—Sean $A=[a_j^i]$ y $B=[b_j^i]$ dos matrices (m, n) sobre K. Se define la suma $A+B$ por la regla

$$A+B=[a_j^i+b_j^i],$$

que equivale a sumar los elementos de lugares homólogos en A y B.

EJEMPLO:

$$\begin{bmatrix} 1 & 2 & 5 \\ -1 & 0 & 8 \end{bmatrix}+\begin{bmatrix} -1 & 2 & 3 \\ 0 & 1 & -4 \end{bmatrix}=\begin{bmatrix} 0 & 4 & 8 \\ -1 & 1 & 4 \end{bmatrix}$$

Es evidente que el conjunto de las matrices (m, n) queda dotado de una estructura de grupo aditivo. El elemento neutro es la matriz nula $O=[0]$, cuyos elementos son todos iguales a cero; la opuesta de $A=[a_j^i]$ es $-A=[(-a_j^i)]$.

TEOREMA.—*La traspuesta de una suma es la suma de las traspuestas.*

$${}^t(A+B)={}^tA+{}^tB.$$

5. Espacio vectorial de las matrices (m, n). Multiplicación por un escalar.—Se forma el producto de A por $\lambda \in K$, multiplicando por λ cada elemento de A; p. ej.,

$$6\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}=\begin{bmatrix} 6 & 12 \\ 0 & 18 \end{bmatrix}$$

Es inmediato ver que el elemento unidad de K es neutro, y que

$$\lambda(\mu A)=(\lambda\mu)A; \qquad (\lambda+\mu)A=\lambda A+\mu A; \qquad \lambda(A+B)=\lambda A+\lambda B.$$

Las dos leyes de composición precedentes hacen del conjunto de las matrices de dimensiones dadas un espacio vectorial. Además, una matriz (m, n) está definida por mn elementos ordenados de K, puesto que los elementos de una fila están ordenados y también lo está el conjunto de las filas. El espacio obtenido es, por tanto, isomorfo a K^{mn} y comprende mn dimensiones.

TEOREMA.—*Las matrices* (m, n) *sobre un cuerpo* K *forman sobre este cuerpo un espacio vectorial de dimensión* mn.

Base canónica.—La base canónica está formada por el conjunto de las mn matrices del tipo E_j^i que tienen nulos todos sus elementos a excepción de uno solo, situado en la intersección de la fila i-ésima con la columna de lugar j, y que es la unidad de K. Una matriz (m, n) A se escribe por medio de esta base:

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_j^i E_j^i.$$

EJEMPLO.—Las matrices cuadradas de segundo orden forman un espacio vectorial de cuatro dimensiones. Tenemos, p. ej.,

$$\begin{bmatrix} 1 & 5 \\ -2 & 8 \end{bmatrix} = 1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 5\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - 2\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 8\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= 1E_1^1 + 5E_2^1 - 2E_1^2 + 8E_2^2.$$

Matriz antisimétrica.—Una matriz se llama antisimétrica cuando es opuesta de su traspuesta: ${}^tA = -A$, lo que supone cuadrada la matriz. Tenemos $a_j^i = -a_i^j$ y, por consiguiente, $a_i^i = 0$.

TEOREMA.—*Toda matriz cuadrada es la suma de una matriz simétrica y de otra antisimétrica; esta descomposición es única.*

Admitiendo una descomposición $A = S + T$, con S simétrica y T antisimétrica, se tiene que ${}^tA = S - T$; por consiguiente,

$$S = \frac{1}{2}(A + {}^tA), \qquad T = \frac{1}{2}(A - {}^tA).$$

Observemos que el teorema pierde todo significado si el cuerpo K es de característica 2 (véase ejercicio 5, a continuación).

EJERCICIOS:

1. Determínense dos matrices, X e Y, cuadradas y de orden 2, que verifiquen al sistema lineal

$$2X-5Y=\begin{bmatrix}1&-2\\0&1\end{bmatrix}, \qquad -X+3Y=\begin{bmatrix}2&1\\3&0\end{bmatrix}.$$

2. Sean dos conjuntos E y F, dos relaciones $\mathcal{R}$ y $\mathcal{S}$ y sus matrices de verdad A y B.

a) Se designa por $\mathcal{R}\cup\mathcal{S}$ cualquier relación que sea verdad si lo es una al menos de las relaciones. Demuéstrese que se forma su matriz «sumando» las matrices A y B de acuerdo con la regla ordinaria, pero teniendo en cuenta, además, las leyes

$$0\cup 0=0, \qquad 1\cup 0=0\cup 1=1\cup 1=1.$$

b) Se designa por $\mathcal{R}\cap\mathcal{S}$ aquella relación que es verdad si, y solo si, las dos lo son. Fórmese su matriz inspirándose en *a)*.

3. Estúdiese la dependencia de las tres matrices

$$A=\begin{bmatrix}2&-2\\4&-6\end{bmatrix}, \qquad B=\begin{bmatrix}1&-4\\5&3\end{bmatrix}, \qquad C=\begin{bmatrix}0&-4\\-4&-8\end{bmatrix}$$

4. Determínense λ y μ de manera que las matrices

$$\begin{bmatrix}2&1\\5&3\end{bmatrix}, \qquad \begin{bmatrix}5&-3\\2&-1\end{bmatrix}, \qquad \begin{bmatrix}-11&\lambda\\4&\mu\end{bmatrix}$$

sean dependientes. Relación de dependencia.

5. Demuéstrese que sobre el cuerpo de los enteros módulo 2 toda matriz antisimétrica es también simétrica.

6. Subespacios vectoriales.—En el espacio de las matrices (m, n) se definen numerosos subespacios. Indiquemos algunos ejemplos.

Matrices diagonales.—Se llama así una matriz cuadrada en la que todos los elementos a_i^j que corresponden a índices i y j diferentes son nulos. Los términos de la forma a_i^i son cualesquiera y se colocan sobre la *diagonal principal*. Basta un solo índice para denotar tal matriz, que se escribe

$$\text{diag } (a_1, a_2, \ldots, a_n).$$

El conjunto de las matrices diagonales cuadradas de orden n constituye un subespacio de n dimensiones.

Matrices triangulares.—Se denomina así una matriz cuadrada en la que todos los elementos situados por encima (o por debajo) de la diagonal

principal son nulos. Para el orden n, constituyen un subespacio de $n(n+1)/2$ dimensiones.

EJERCICIOS:

1. Demuéstrese que las matrices simétricas de orden n constituyen un subespacio vectorial. ¿Cuál es su dimensión? Indíquese otro subespacio de matrices, isomorfo del precedente.

2. Las mismas cuestiones para el conjunto de las matrices antisimétricas.

3. Se considera el conjunto de las matrices de la forma

$$A=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

siendo a y b reales. Pruébese que constituye un subespacio del conjunto de las matrices de orden 2. Dése una base del mismo.

4. Sobre el cuerpo de los reales se considera el conjunto de las matrices de la forma

$$A=\begin{bmatrix} a & b \\ -qb & a+pb \end{bmatrix}$$

donde p y q son dados, y a y b, arbitrarios. ¿Se obtiene un subespacio vectorial? En caso afirmativo, dése una base del mismo.

RANGO DE UNA MATRIZ. APLICACION A LOS SISTEMAS LINEALES

7. **Teorema fundamental.**—Sea una matriz (m, n), A, sobre un cuerpo K:

$$A=\begin{bmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \vdots & & & \vdots \\ a_m^1 & a_m^2 & \cdots & a_m^n \end{bmatrix}$$

Los vectores columna $\mathbf{A}^1$, $\mathbf{A}^2$, ..., $\mathbf{A}^n$ forman n vectores del espacio K^m; son de rango r. Los vectores fila $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_m$ constituyen m vectores de K^n; sea r' su rango.

TEOREMA.—*El rango de los vectores fila de una matriz es igual al rango de los vectores columna.*

Se puede suponer siempre que los r primeros vectores columna son independientes, porque una sustitución sobre las columnas, aunque cam-

bia los vectores fila (sustitución sobre sus coordenadas), sabemos que conserva su rango (véase pág. 135). Las columnas de orden $r+1$, $r+2$, ..., n son así combinaciones lineales de las r primeras, de donde las $n-r$ igualdades

$$\mathbf{A}^{r+\alpha} = \lambda_1^\alpha \mathbf{A}^1 + \lambda_2^\alpha \mathbf{A}^2 + \ldots + \lambda_r^\alpha \mathbf{A}^r \qquad (\alpha \text{ de I a } n-r).$$

Las coordenadas de orden i de los vectores columna forman una fila de la matriz, y las relaciones precedentes permiten expresar los $n-r$ últimos elementos de una fila como funciones lineales de los r primeros, *siendo independientes los coeficientes de la fila considerada:*

$$a_i^{r+\alpha} = \lambda_1^\alpha a_i^1 + \lambda_2^\alpha a_i^2 + \ldots + \lambda_r^\alpha a_i^r.$$

Demostremos que esto permite una reducción del número de generadores del subespacio fila. Si $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ son los vectores de la base canónica de K^n, tenemos:

$$\mathbf{A}_i = \sum_1^n a_i^p \mathbf{e}_p$$

donde se pueden reemplazar los $n-r$ últimos coeficientes por las expresiones anteriores; se llega así a

$$\mathbf{A}_i = \sum_1^r a_i^p \mathbf{e}_p + \sum_{\alpha=1}^{\alpha=n-r} (\lambda_1^\alpha a_i^1 + \ldots + \lambda_r^\alpha a_i^r) \mathbf{e}_{r+\alpha}.$$

La independencia de los coeficientes λ con relación al orden de la fila permite un reagrupamiento interesante de los vectores. Agrupando los términos que contienen a_i^p, se escribe:

$$\mathbf{A}_i = \sum_1^r a_i^p \mathbf{e}'_p,$$

haciendo, para $1 \leqslant p \leqslant r$,

$$\mathbf{e}'_p = \mathbf{e}_p + \sum_{\alpha=1}^{\alpha=n-r} \lambda_p^\alpha \mathbf{e}_{r+\alpha}.$$

Los m vectores fila son, por consiguiente, combinaciones lineales de los r vectores $\mathbf{e}'_p$ de K^n. Su rango r' es inferior o igual al número r

de los generadores: $r' \leqslant r$. Pero el razonamiento anterior, en el que los papeles de las filas y de las columnas no son simétricos, puede repetirse cambiando estos papeles. Se concluye que $r \leqslant r'$; la compatibilidad de las dos desigualdades exige $r = r'$.

Definición.— El rango común a los vectores fila y a los vectores columna se llama *rango* o *característica* de la matriz. En una matriz de rango r, todo conjunto de $r+1$ filas (o $r+1$ columnas) es dependiente, pero es posible hallar un sistema al menos de r filas y otro de r columnas independientes. Muchas veces es cómodo suponer que las r primeras filas y columnas satisfacen estas condiciones; para reducirlo siempre a este caso, basta efectuar sustituciones de columnas o de filas, lo cual conserva el rango. Es necesario también comprobar si esto es admisible para la cuestión que se esté estudiando.

Para hallar prácticamente el rango de una matriz dada, se determina, p. ej., el rango de las filas por medio del método estudiado en la página 140.

EJERCICIOS:

1. Determínese el rango de cada una de las siguientes matrices:

$$\begin{bmatrix} 5 & 8 & 2 & 3 \\ 2 & 2 & 2 & -1 \\ 20 & 10 & 2 & 5 \\ -2 & 0 & -1 & 1 \\ 8 & -6 & -3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & -2 & -1 \\ 2 & 5 & -2 & 1 \\ 1 & 1 & 6 & 13 \end{bmatrix}, \quad \begin{bmatrix} 2 & -5 & 1 & 4 \\ 3 & 2 & 6 & 5 \\ -1 & 12 & 4 & -3 \\ 11 & -37 & 1 & 23 \end{bmatrix}$$

2. Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, n vectores de un espacio vectorial. Les asociamos p combinaciones lineales

$$\mathbf{w}_i = \sum_1^n a_{i\alpha}\mathbf{v}_\alpha \qquad (i \text{ de } 1 \text{ a } p).$$

Si los vectores $\mathbf{v}_i$ son independientes, demuéstrese que el rango de los vectores $\mathbf{w}$ es igual al de la matriz $\mathrm{A} = [a_{i\alpha}]$.

3. Demuéstrese que toda matriz cuadrada de rango 1 y orden 3 tiene la forma

$$\begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix}.$$

4. Se considera la matriz $\mathrm{C} = [\mathrm{A} \,;\, \mathrm{B}]$ obtenida yuxtaponiendo una matriz (m, n), A, y otra matriz (m, p), B. Demuéstrese que rango C $\leqslant$ rango de A + rango de B.

8. Matriz regular.—Una matriz (m, n), A, se llama regular cuando sus m vectores fila son independientes, así como también sus n vectores columna. Si r es el rango, se debe tener $m=r$ y $n=r$, de donde $m=n$; por tanto, solo las matrices cuadradas pueden ser regulares. Una matriz no regular se llama *singular*.

Matrices extraídas de una matriz dada.—Dada una matriz, A, toda matriz deducida de A por supresión de cierto número de filas y de cierto número de columnas (sin cambiar el orden de las restantes) se dice extraída de A.

TEOREMA.—*El rango de una matriz* (m, n), A, *es igual al orden más elevado de las matrices cuadradas regulares extraídas de* A.

Supongamos que A contiene al menos una matriz cuadrada regular, B, de orden s, siendo singulares todas las matrices cuadradas de orden superior a s que se pueden extraer de A; probemos que $s=r$.

Observemos, en primer lugar, que las s filas de A que intervienen para formar B son independientes, porque la dependencia de las filas de A implica la de las filas de B, y B no sería regular. De donde resulta, $r \geqslant s$.

Demostremos por último que $s+1$ filas de A son dependientes. Si un conjunto de $s+1$ filas fuera independiente, la matriz C, extraída de A y formada por dichas filas, sería de rango $s+1$; tendría, por tanto, $s+1$ columnas independientes, y con ellas se podría extraer de C (por consiguiente, de A) una matriz cuadrada regular de orden $s+1$, lo que es contrario a la elección de s. Las $s+1$ filas son, pues, dependientes, de donde se deduce que $r < s+1$. Las dos condiciones obtenidas demuestran que $r=s$.

OBSERVACIÓN.—El teorema anterior proporciona un nuevo método para determinar el rango, estudiando la regularidad de las matrices cuadradas extraídas de A, lo que podrá hacerse con ayuda de los determinantes. Una matriz cuadrada regular de orden máximo extraída de A se llama *matriz principal*.

9. Sistema lineal. Interpretación vectorial.—Consideremos m formas lineales con n variables, $x_1, x_2, \ldots, x_n$, y sea

$$f_i = a_i^1 x_1 + a_i^2 x_2 + \ldots + a_i^n x_n \qquad (i \text{ de } 1 \text{ a } m).$$

Elegidos los coeficientes en un cuerpo K, se trata de hallar en el mismo cuerpo qué valores es necesario atribuir a las variables, llama-

das incógnitas, para que las formas tomen m valores dados $b_1, b_2, \ldots, b_m$ de K. El sistema más general de m ecuaciones con n incógnitas es de la forma:

$$\left\{\begin{array}{l} a_1^1 x_1 + a_1^2 x_2 + \ldots + a_1^n x_n = b_1, \\ a_2^1 x_1 + a_2^2 x_2 + \ldots + a_2^n x_n = b_2, \\ \quad\vdots \qquad\quad \vdots \qquad\qquad \vdots \\ a_m^1 x_1 + a_m^2 x_2 + \ldots + a_m^n x_n = b_m. \end{array}\right. \qquad [I]$$

De manera abreviada, se escribe:

$$\sum_{\alpha=1}^{\alpha=n} a_i^\alpha x_\alpha = b_i.$$

El índice i corresponde al lugar que ocupa la ecuación en el sistema (i varía de 1 a m), y el índice mudo α, a la sumación de los términos en una misma ecuación (α varía de 1 a n).

A este sistema se asocian dos matrices: la matriz A de los coeficientes de las incógnitas, de la que designaremos por $\mathbf{A}^1$, $\mathbf{A}^2$, ..., $\mathbf{A}^n$ los vectores columna, y la matriz completa, que contiene una nueva columna formada por los elementos del segundo miembro:

$$A = [A^1\, A^2 \ldots A^n], \qquad [AB] = [A^1\, A^2 \ldots A^n\, B].$$

Los vectores $\mathbf{A}^1$, $\mathbf{A}^2$, ..., $\mathbf{A}^n$, $\mathbf{B}$ pertenecen todos al espacio K^m y el sistema [I] se resume en una igualdad vectorial única

$$x_1\mathbf{A}^1 + x_2\mathbf{A}^2 + \ldots + x_n\mathbf{A}^n = \mathbf{B} \qquad [II]$$

cuya interpretación es inmediata: el cálculo de los sistemas de elementos $(x_1, x_2, \ldots, x_n)$ equivale a estudiar todas las descomposiciones posibles de $\mathbf{B}$ según los vectores $\mathbf{A}^1$, $\mathbf{A}^2$, ..., $\mathbf{A}^n$.

10. Condición de posibilidad. **Teorema de Kronecker.**—*Es condición necesaria y suficiente, para que un sistema de m ecuaciones con n incógnitas admita al menos una solución, que el rango de la matriz de los coeficientes sea igual al rango de la matriz completa del sistema.*

La igualdad [II] prueba que, si existe una solución, $\mathbf{B}$ pertenece al subespacio de los vectores columna de la matriz A; luego el rango de las columnas (es decir, el rango) de [A B] es el mismo que el de [A].

Recíprocamente, si el rango es el mismo y si $\mathbf{A}^1, \mathbf{A}^2, \ldots, \mathbf{A}^r$ designan los vectores columna independientes de A, **B** pertenece necesariamente al subespacio engendrado por estas columnas; en otro caso, el rango de [A|B] sería $r+1$. Se puede descomponer **B** de manera única según los vectores independientes $\mathbf{A}^1, \mathbf{A}^2, \ldots, \mathbf{A}^r$, lo que da los valores de las incógnitas $x_1, x_2, \ldots, x_r$, y, eligiendo $x_{r+1}=x_{r+2}=\ldots=x_n=0$, queda determinada una solución del sistema.

Conjunto de las soluciones.—Precisemos cómo se obtienen todas las soluciones. Designemos como antes por $\mathbf{A}^1, \mathbf{A}^2, \ldots, \mathbf{A}^r$ los vectores columna independientes de A. Se supone que estos son los r primeros, lo que siempre puede lograrse cambiando en caso necesario el orden de las incógnitas; $x_1, x_2, \ldots, x_r$ se llaman *incógnitas principales* [1]. Escribamos la ecuación [II] poniendo en evidencia los vectores independientes:

$$x_1\mathbf{A}^1 + x_2\mathbf{A}^2 + \ldots + x_r\mathbf{A}^r = \mathbf{B} - x_{r+1}\mathbf{A}^{r+1} - \ldots - x_n\mathbf{A}^n.$$

Cualquiera que sea la elección de las incógnitas no principales, $x_{r+1}, \ldots, x_n$, el vector del segundo miembro pertenece al subespacio columna de A y se descompone de manera única según $\mathbf{A}^1, \mathbf{A}^2, \ldots, \mathbf{A}^r$. Recíprocamente, toda solución del sistema puede obtenerse de esta manera; de donde resultan las tres posibilidades siguientes:

1.ª la condición de Kronecker no se verifica: no existe solución;

2.ª se cumple dicha condición y el rango r es igual al número n de incógnitas: no hay incógnitas no principales y el sistema admite solución única;

3.ª se verifica la condición y $r < n$: existe una infinidad de soluciones.

Sistemas de Cramer.—El caso más importante se obtiene cuando existen tantas ecuaciones como incógnitas: $m=n$. Si la matriz cuadrada A es regular ($r=n$) el teorema de Kronecker se verifica con certeza, porque el rango de [A|B] no puede superar a n. Nos hallamos, pues, en el segundo caso y el sistema admite solución única. Se llama sistema de Cramer.

11. La condición de Kronecker.—Vamos a expresar que las dos matrices A y [A|B] tienen el mismo rango r. Se puede suponer, en caso necesario se cambia la ordenación de las ecuaciones, que la ma-

[1] La elección no es, en general, única.

triz de orden r, A_r, extraída de A, y que está formada por los elementos de las r primeras filas y de las r primeras columnas, es principal. Le asociamos ciertas matrices cuadradas de orden $r+1$, que se extraen de $[A|B]$ suprimiendo las columnas de orden $r+1$, $r+2$, ..., n (pero conservando la última) y suprimiendo las $m-r$ últimas filas, salvo una cualquiera de ellas. Hay $m-r$ matrices de esta especie según la elección de la fila de orden superior a r que se haya conservado. Decimos que las $m-r$ matrices así formadas, que comprenden A_r ampliada con una fila y una columna suplementarias, son las *matrices orladas* de A_r.

TEOREMA.—*Para que la matriz completa del sistema tenga el mismo rango que la matriz de los coeficientes es necesario y suficiente que las matrices orladas de la matriz principal sean singulares.*

La condición es necesaria porque, si $[A|B]$ es de rango r, toda matriz cuadrada de orden $r+1$ extraída de ella es singular. Demostremos, recíprocamente, que si las matrices obtenidas orlando A, son singulares, $[A|B]$ es de rango r, ya que sus $m-r$ últimas filas son combinaciones lineales de las r primeras.

Consideremos, en efecto, una matriz C extraída de $[A|B]$ suprimiendo las $m-r$ últimas filas, salvo una, y demostremos que C es de rango r. Estudiemos sus vectores columna: los de órdenes $r+1$, $r+2$, ..., n son combinaciones lineales de los r primeros, porque esto ya sucedía en A; por otra parte, la matriz orlada de A_r y correspondiente a C (con la misma última fila) es singular, lo que implica que la última columna de C es combinación lineal de las r primeras (estas son independientes); el subespacio columna de C es, pues, de r dimensiones y, como sucede lo mismo con el subespacio fila, necesariamente la última fila es combinación lineal de las r primeras, que son independientes.

Si se cumple la misma condición para todas las matrices que resultan al orlar A_r, el rango de $[A|B]$ es necesariamente igual a r, y las condiciones indicadas son suficientes.

OBSERVACIONES.—1.ª Para la discusión efectiva de un sistema de m ecuaciones con n incógnitas, bastará, pues, saber expresar que una matriz cuadrada es o no regular. Un método general se indicará en el capítulo XVI, donde discutiremos un ejemplo.

2.ª La demostración anterior enseña también cómo se puede resolver el sistema: las $m-r$ últimas ecuaciones las verifica todo sistema de números $(x_1, x_2, \ldots, x_n)$ que satisfaga a las r primeras (ecuaciones

principales). Resulta así posible reducir el sistema a sus ecuaciones principales, atribuir valores arbitrarios a las $n-r$ incógnitas no principales y resolver el sistema de Cramer de r ecuaciones con r incógnitas así formado.

PRODUCTO DE MATRICES

12. Producto de dos matrices.—Vamos a definir una nueva ley de composición que, en ciertos casos, permite asociar a dos matrices A y B una tercera matriz C, llamada producto y representada por la notación habitual $C=AB$. Esta operación no es, en general, conmutativa y debemos tener en cuenta el orden de los factores, pero es asociativa y distributiva a la derecha y a la izquierda con relación a la adición, lo que justifica el término *producto*. Hablando del producto AB diremos que A multiplica a B por la izquierda (o que premultiplica a B) y que B multiplica a A por la derecha (o que posmultiplica a A).

Hemos visto (pág. 292) que puede asociarse una matriz a toda sustitución lineal. La definición de producto de dos sustituciones conduce, de manera natural, al producto de dos matrices. Razonemos sobre un ejemplo. Sea la sustitución, definida por las fórmulas

$$\begin{cases} y_1=b_1^1x_1+b_1^2x_2, \\ y_2=b_2^1x_1+b_2^2x_2 \end{cases} \quad \text{de matriz} \quad B=\begin{bmatrix} b_1^1 & b_1^2 \\ b_2^1 & b_2^2 \end{bmatrix}. \qquad [1]$$

A todo par (x_1, x_2) asocia un par (y_1, y_2). Supongamos que se efectúa una nueva sustitución que asocia el par (z_1, z_2) a (y_1, y_2); o sea,

$$\begin{cases} z_1=a_1^1y_1+a_1^2y_2, \\ z_2=a_2^1y_1+a_2^2y_2 \end{cases} \quad \text{de matriz} \quad A=\begin{bmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{bmatrix}. \qquad [2]$$

Cabe considerar como producto de estas sustituciones la que hace pasar del par (x_1, x_2) al (z_1, z_2). Reemplazando y_1 e y_2 deducidos de [1] en las expresiones [2], se obtiene la sustitución producto

$$\begin{cases} z_1=(a_1^1b_1^1+a_1^2b_2^1)x_1+(a_1^1b_1^2+a_1^2b_2^2)x_2, \\ z_2=(a_2^1b_1^1+a_2^2b_2^1)x_1+(a_2^1b_1^2+a_2^2b_2^2)x_2, \end{cases}$$

La matriz C de esta sustitución es, por definición, el producto de las matrices A y B, en este orden:

$$AB=C, \quad \text{o} \quad \begin{bmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{bmatrix} \begin{bmatrix} b_1^1 & b_1^2 \\ b_2^1 & b_2^2 \end{bmatrix} = \begin{bmatrix} a_1^1b_1^1+a_1^2b_2^1 & a_1^1b_1^2+a_1^2b_2^2 \\ a_2^1b_1^1+a_2^2b_2^1 & a_2^1b_1^2+a_2^2b_2^2 \end{bmatrix}.$$

Cada término de C se obtiene multiplicando los elementos de una fila de A por los correspondientes elementos de una columna de B y sumando los resultados. Es la multiplicación fila por columna.

El elemento de la fila i y columna j de C corresponde a la multiplicación de la fila i-ésima de A por la columna de lugar j de B.

Hemos razonado sobre matrices cuadradas muy sencillas, pero la regla se extiende sin dificultad a las matrices rectangulares, con la sola hipótesis de que sea posible la multiplicación de una fila de A por una columna de B. Llegamos, pues, a las siguientes definiciones:

1.ª DEFINICIÓN.—El producto de una matriz fila A de n elementos, por una matriz columna B del mismo número de elementos,

$$\mathrm{A}=[a_1 \quad a_2 \quad \ldots \quad a_n], \qquad \mathrm{B}=\begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \end{bmatrix}$$

es, por definición, una matriz de una fila y una columna $\mathrm{AB}=[c]$, cuyo único elemento c es

$$c=a_1b^1+a_2b^2+\ldots+a_nb^n=\sum_\alpha a_\alpha b^\alpha.$$

Pasemos al caso general: A, de dimensiones (m, n), se considera formada por m matrices fila de n elementos, y B, de dimensiones (n, p), está formada por p matrices columna de n elementos. La restricción sobre las dimensiones (m, n), (n, p) permite la multiplicación de una fila cualquiera de A por una columna cualquiera de B según la fórmula anterior.

2.ª DEFINICIÓN (CASO GENERAL).—El producto de una matriz (m, n), A, por una matriz (n, p), B, es la matriz C cuyo elemento de la fila i y columna j es el producto de la i-ésima matriz fila de A por la j-ésima matriz columna de B.

El término general c_i^j de la fila i y columna j de C es, por consiguiente,

$$c_i^j=\sum_\alpha a_i^\alpha b_\alpha^j \qquad (\alpha \text{ de 1 a } n).$$

De manera general, se tiene:

$$\begin{bmatrix} a_1^1 & \cdots & a_1^n \\ \vdots & & \vdots \\ a_m^1 & \cdots & a_m^n \end{bmatrix}\begin{bmatrix} b_1^1 & \cdots & b_1^p \\ \vdots & & \vdots \\ b_n^1 & \cdots & b_n^p \end{bmatrix}=\begin{bmatrix} \Sigma a_1^\alpha b_\alpha^1 & \Sigma a_1^\alpha b_\alpha^2 & \cdots & \Sigma a_1^\alpha b_\alpha^p \\ \vdots & \vdots & & \vdots \\ \Sigma a_m^\alpha b_\alpha^1 & \Sigma a_m^\alpha b_\alpha^2 & \cdots & \Sigma a_m^\alpha b_\alpha^p \end{bmatrix}.$$

El producto de una matriz (m, n), A, por una matriz (n, p), B, es una matriz (m, p); A, escrita a la izquierda, impone el número de filas, y B, escrita a la derecha, el número de columnas.

Matrices fila del producto AB.—La definición precedente permite formar término a término el producto AB. Sin embargo, sus filas y columnas tienen con las filas de B o las columnas de A (respectivamente) relaciones muy sencillas que vamos a precisar. Supongamos B descompuesta según sus filas:

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix},$$

y estudiemos la primera fila del producto AB; se escribe

$$[\Sigma a_1^\alpha b_\alpha^1 \quad \Sigma a_1^\alpha b_\alpha^2 \quad \dots \quad \Sigma a_1^\alpha b_\alpha^p].$$

En cada uno de los términos encontramos los mismos coeficientes: los de la primera fila de A. Se puede, pues, descomponer esta matriz en suma de n matrices fila:

$$a_1^1[b_1^1 \ b_1^2 \ \dots \ b_1^p] + a_1^2[b_2^1 \ b_2^2 \ \dots \ b_2^p] + \dots + a_1^n[b_n^1 \ b_n^2 \dots b_n^p]$$

o

$$a_1^1 B_1 + a_1^2 B_2 + \dots + a_1^n B_n.$$

Se obtiene un resultado análogo para una fila cualquiera del producto; de todo lo cual se deduce que:

$$\begin{bmatrix} a_1^1 & \dots & a_1^n \\ a_2^1 & \dots & a_2^n \\ \vdots & & \vdots \\ a_m^1 & \dots & a_m^n \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} \Sigma a_1^\alpha B_\alpha \\ \Sigma a_2^\alpha B_\alpha \\ \vdots \\ \Sigma a_m^\alpha B_\alpha \end{bmatrix}.$$

TEOREMA 1.—*El producto de las matrices* A *y* B *es una matriz* AB *que admite por i-ésima fila la combinación lineal de las filas de* B *cuyos coeficientes son los términos del mismo orden de la i-ésima fila de* A.

Recíprocamente, toda matriz C cuyas filas sean combinaciones lineales de las filas de una matriz B se obtiene multiplicando B a la

izquierda por una matriz A de la que se conocen de manera inmediata las filas.

Estudiemos ahora las columnas de AB; la de orden j, p. ej., es una combinación lineal de las columnas de A:

$$\begin{bmatrix} \Sigma a_1^\alpha b_\alpha^j \\ \Sigma a_2^\alpha b_\alpha^j \\ \vdots \\ \Sigma a_m^\alpha b_\alpha^j \end{bmatrix} = b_1^j \begin{bmatrix} a_1^1 \\ a_2^1 \\ \vdots \\ a_m^1 \end{bmatrix} + b_2^j \begin{bmatrix} a_1^2 \\ a_2^2 \\ \vdots \\ a_m^2 \end{bmatrix} + \ldots + b_n^j \begin{bmatrix} a_1^n \\ a_2^n \\ \vdots \\ a_m^n \end{bmatrix}.$$

Se ha establecido así el teorema siguiente:

TEOREMA 2.—*El producto de las matrices* A *y* B *es una matriz* AB *que admite por j-ésima columna la combinación lineal de las columnas de* A *cuyos coeficientes son los términos correspondientes de la j-ésima columna de* B.

Resulta de dos propiedades anteriores que el rango del producto AB es inferior o igual a los rangos de A y de B.

En efecto: el subespacio fila de AB está incluido en el subespacio fila de B, y el subespacio columna de AB queda incluido en el subespacio columna de A.

13. Propiedades del producto. **Asociatividad.**—Es la propiedad fundamental de la operación. Se pueden multiplicar varias matrices A, B, C, en este orden, con la condición de que el número de columnas de cada una de ellas sea igual al de filas de la siguiente; así, p. ej.,

$$(m \text{ fi.};\ n \text{ col.}) \times (n \text{ fi.};\ p \text{ col.}) \times (p \text{ fi.};\ q \text{ col.}) = (m \text{ fi.};\ q \text{ col.}).$$

Demostremos que el producto es asociativo: $(AB)C = A(BC)$.

Cabe considerar que cada uno de los productos precedentes efectúa combinaciones lineales con las filas de C. Pongamos

$$A = [a_i^j], \qquad B = [b_k^j], \qquad C = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_p \end{bmatrix}$$

Se tiene, en primer lugar,

$$BC = \begin{bmatrix} b_1^1 & \ldots & b_1^p \\ \vdots & & \vdots \\ b_n^1 & \ldots & b_n^p \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_p \end{bmatrix} = \begin{bmatrix} b_1^1 C_1 + b_1^2 C_2 + \ldots + b_1^p C_p \\ b_2^1 C_1 + b_2^2 C_2 + \ldots + b_2^p C_p \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_n^1 C_1 + b_n^2 C_2 + \ldots + b_n^p C_p \end{bmatrix}$$

La multiplicación a la izquierda por A equivale a realizar combinaciones lineales con las filas de BC. Agrupando los términos en $C_1, C_2, \ldots, C_p$, se obtiene:

$$A(BC) = \begin{bmatrix} (\Sigma a_1^\alpha b_\alpha^1)C_1 + (\Sigma a_1^\alpha b_\alpha^2)C_2 + \ldots + (\Sigma a_1^\alpha b_\alpha^p)C_p \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (\Sigma a_m^\alpha b_\alpha^1)C_1 + (\Sigma a_m^\alpha b_\alpha^2)C_2 + \ldots + (\Sigma a_m^\alpha b_\alpha^p)C_p \end{bmatrix} \quad (\alpha \text{ de } 1 \text{ a } n).$$

Se ve que en la matriz obtenida los coeficientes de $C_1, C_2, \ldots, C_p$, de una fila determinada, son los elementos del mismo lugar de una fila del producto AB. Queda, pues, demostrado, que $A(BC) = (AB)C$.

Distributividad.—La distributividad a la derecha con relación a la adición,

$$A(B_1 + B_2) = AB_1 + AB_2$$

resulta fácilmente del teorema 2 (pág. 308), porque las operaciones en ambos miembros efectúan las mismas combinaciones lineales de las columnas de A.

La distributividad a la izquierda,

$$(A_1 + A_2)B = A_1B + A_2B$$

resulta de manera análoga del teorema 1.

Señalemos, por último, que respecto de la multiplicación por un escalar $\lambda \in K$, se tiene:

$$(\lambda A)B = A(\lambda B) = \lambda AB.$$

Elemento neutro.—Tratemos de hallar un elemento neutro para la multiplicación a la izquierda de matrices (m, n). La matriz buscada debe ser cuadrada y de orden m; según el teorema 1 (pág. 307), cada fila no debe contener más que un solo elemento no nulo e igual a 1, el del mismo orden que la fila. Se obtiene así una matriz diagonal I_m, de la que todos los elementos no nulos sobre la diagonal son iguales a 1:

$$I_m = \text{diag}\,(1, 1, \ldots, 1).$$

Igualmente, la matriz I_n, construida de manera análoga pero de orden n, es neutra para la multiplicación a la derecha de matrices (m, n). En particular, I_n es neutra a la derecha y a la izquierda sobre el conjunto de las matrices cuadradas de orden n.

La igualdad $AI_n = A$ puede utilizarse para escribir la matriz A realizando combinaciones lineales determinadas de las n filas de una matriz cualquiera; A se obtiene efectuando sobre las filas de I_n las combinaciones buscadas.

No conmutatividad.—En general, no cabe hablar de conmutatividad, ya que si es A una matriz (m, n) y B una matriz (n, p), el producto BA es solo posible cuando $m = p$; pero, en tal caso, AB es cuadrada de orden m, en tanto que BA es cuadrada de orden n. No se obtienen dos productos de igual dimensión más que multiplicando matrices cuadradas del mismo orden. Sobre un ejemplo el lector comprobará que en general $AB \neq BA$, en razón de los distintos papeles que desempeñan las filas y las columnas en la multiplicación.

Traspuesta de un producto.—Estudiemos la matriz traspuesta del producto AB. Según el teorema 1 (pág. 307),

$$AB = \begin{bmatrix} a_1^1 B_1 + \ldots + a_1^n B_n \\ a_2^1 B_1 + \ldots + a_2^n B_n \\ \ldots\ldots\ldots\ldots\ldots \\ a_m^1 B_1 + \ldots + a_m^n B_n \end{bmatrix}.$$

Llamando B^i la matriz columna traspuesta de la fila B_i y trasponiendo las filas del producto AB, se obtiene

$${}^t(AB) = [\Sigma a_1^\alpha B^\alpha,\ \Sigma a_2^\alpha B^\alpha,\ \ldots,\ \Sigma a_m^\alpha B^\alpha]. \qquad (\alpha \text{ de } 1 \text{ a } n).$$

Puesto que las B^α son las columnas de ${}^tB$, se reconoce en esto (teorema 2, pág. 308) la multiplicación a la derecha de ${}^tB$ por cierta matriz. De manera precisa,

$${}^t(AB) = [B^1\ B^2\ \ldots\ B^n] \begin{bmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & & a_m^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & & a_m^n \end{bmatrix}.$$

Como la última matriz es la traspuesta de A, resulta:

$${}^t(AB) = {}^tB\,{}^tA.$$

TEOREMA.—*Se obtiene la traspuesta de un producto multiplicando en orden inverso las traspuestas de sus diferentes factores.*

OBSERVACIONES.—1.ª Las propiedades precedentes son ciertas sobre

un cuerpo cualquiera K; en la práctica se utiliza, sobre todo, R o C. En este último caso (cuerpo C) se puede asociar a toda matriz (m, n), $A=[a_i^j]$, su *conjugada*, denotada por $\bar{A}$; es la matriz (m, n) obtenida reemplazando en A cada término por su conjugado: $\bar{A}=[\bar{a}_i^j]$.

La conjugación obedece a las reglas siguientes:

$$\bar{\bar{A}}=A, \qquad \overline{A+B}=\bar{A}+\bar{B}, \qquad \overline{AB}=\bar{A}\bar{B}.$$

La comprobación de estas fórmulas no plantea dificultad; para la última, observemos que

$$c_i^j=\Sigma a_i^a b_a^j \qquad \text{implica} \qquad \bar{c}_i^j=\Sigma \bar{a}_i^a \bar{b}_a^j.$$

2.ª Sobre C se hace también corresponder a A su *matriz asociada:* es la conjugada de la traspuesta de A, es decir, la matriz ${}^t\bar{A}$. Se denota por A^*, comprobándose fácilmente las siguientes fórmulas:

$$(A^*)^*=A, \qquad (A+B)^*=A^*+B^*, \qquad (AB)^*=B^*A^*.$$

Una matriz idéntica a su asociada se llama *hermítica*. Una matriz hermítica real es simétrica.

EJERCICIOS:

1. Sean $A=\begin{bmatrix}1 & 2\\ 3 & 5\end{bmatrix}$ y $B=\begin{bmatrix}18 & 19\\ -16 & 13\end{bmatrix}$. Compárense AB y BA.

2. Sea $A=\begin{bmatrix}1 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 1\end{bmatrix}$. Compárese $A^2=AA$ con A.

3. Sean $A=\begin{bmatrix}a & a'\\ b & b'\\ c & c'\end{bmatrix}$ y $B=\begin{bmatrix}\alpha & \beta & \gamma\\ \alpha' & \beta' & \gamma'\end{bmatrix}$. Compárense AB y BA.

4. Extiéndase por recurrencia la regla de trasposición de un producto al caso de n factores.

5. Si A es una matriz (m, n), ¿qué podemos decir de $A\,{}^tA$ y de ${}^tAA$?

6. Si B es antisimétrica, ¿qué cabe decir de ${}^tABA$?

7. Si A es antisimétrica, ¿qué se puede decir de A^2?

8. Demuéstrese que, si los productos AB y BA de dos matrices rectangulares son posibles, las sumas de los elementos diagonales de AB y BA son iguales. Generalícese la propiedad al caso de tres factores: ABC, BCA, CAB.

9. Se consideran dos matrices cuadradas, J_m y J_n, de órdenes respectivos m y n y de las que todos sus elementos son 1. ¿Qué representa J_mAJ_n si A es una matriz (m, n)?

10. Para que B *conmute* con A (es decir, para que AB=BA) definida por

$$A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{es suficiente que} \quad B=\begin{bmatrix} \alpha & \beta & \gamma & \delta \\ 0 & \alpha & \beta & \gamma \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & 0 & \alpha \end{bmatrix}.$$

Demuéstrese que esta condición es necesaria.

11. Demuéstrese que las matrices de bases canónicas se multiplican según la regla: $E_j^i F_l^k = \delta_j^k G_l^i$, donde E_j^i es del tipo (m, n), F_l^k del tipo (n, p) y G_l^i del tipo (m, p) (el símbolo δ_j^k está definido en la pág. 137). Para $m=n$, estúdiese el caso particular $E_j^i E_l^k$.

12. Utilícense los resultados anteriores para deducir de nuevo la fórmula general de la multiplicación, descomponiendo las matrices según sus bases canónicas y aplicando la distributividad.

13. Para que una matriz cuadrada A de orden n, conmute con una matriz B cualquiera del mismo orden, es necesario y suficiente que conmute con las n^2 matrices de la base canónica. Dedúzcase de lo anterior la forma muy particular de A.

14. Demuéstrese que sobre el cuerpo de los complejos toda matriz A puede escribirse en la forma $A=H+iK$, con H y K hermíticas.

15. (Notaciones del ejemplo 4, párrafo 1, y del ejercicio 2, párrafo 5.) Sea X una parte de E. Se asocia a X la matriz fila $\mathfrak{X}$ obtenida escribiendo 1 sobre la r-ésima columna si $a_r \in X$, y 0 en caso contrario. Designemos, de otra parte, por $\mathcal{R}(x)$ el conjunto de los $y \in F$ tales que $x\mathcal{R}y$ sea verdad; y por $\mathcal{R}(X)$ la reunión de todos los $\mathcal{R}(x)$ cuando x describe X.

Pruébese que se obtiene la matriz asociada a $\mathcal{R}(X)$ multiplicando $\mathfrak{X}$ por la matriz de verdad de $\mathcal{R}$, considerando $\cup$ como la adición.

16. Las mismas notaciones con un tercer conjunto finito $G=\{c_1, c_2, \ldots, c_q\}$ y una relación $\mathcal{S}$ entre los b_j y c_k. Si existe un b_λ de F tal que $a_i \mathcal{R} b_\lambda$ y $b_\lambda \mathcal{S} c_k$ sean ciertas las dos, escribimos $a_i \mathcal{R}\mathcal{S} c_k$ y decimos que $\mathcal{R}\mathcal{S}$ es la relación compuesta de $\mathcal{R}$ y $\mathcal{S}$.

Demuéstrese que la matriz de verdad de $\mathcal{R}\mathcal{S}$ es el producto de las matrices de verdad de $\mathcal{R}$ y de $\mathcal{S}$, efectuando las sumas como en el ejercicio anterior.

14. Ejemplos de fórmulas matriciales.

El empleo del producto matricial permite dar forma condensada a ciertas expresiones algebraicas importantes.

1. Sustitución lineal.—El producto de una matriz (m, n) A por una matriz columna X con n filas proporciona una matriz columna de la que las filas son formas lineales con relación a los elementos de X:

$$\begin{bmatrix} a_1^1 & a_1^2 & \ldots & a_1^n \\ \vdots & \vdots & & \vdots \\ a_m^1 & a_m^2 & \ldots & a_m^n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1^1 x_1 + a_1^2 x_2 + \ldots + a_1^n x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_m^1 x_1 + a_m^2 x_2 + \ldots + a_m^n x_n \end{bmatrix}.$$

Las filas de la matriz del segundo miembro son las expresiones de los m números $(y_1, y_2, \ldots, y_m)$ que corresponden a $(x_1, x_2, \ldots, x_n)$ en la sustitución lineal de matriz A (véase pág. 292). La sustitución se escribe, pues, en forma matricial:

$$Y = AX, \qquad \text{con} \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \qquad \text{y} \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Si se efectúa una sustitución de matriz B sobre $(y_1, y_2, \ldots, y_n)$, se obtiene igualmente

$$Z = BY = B(AX) = (BA)X.$$

Se comprueba que la matriz asociada al producto de las sustituciones de matrices A y B es BA, resultado fácil de prever, puesto que constituye la base de la definición de producto de matrices (véase página 305).

2. Sistema lineal.—En la fórmula $Y = AX$ supongamos conocida Y e igual a una matriz columna B. Cabe considerar la igualdad

$$AX = B$$

como una ecuación matricial en la que la incógnita es la matriz columna X. Según las fórmulas anteriores, esta única ecuación condensa en escritura matricial todo un sistema lineal de m ecuaciones con n incógnitas.

3. Forma bilineal.—Sean X e Y dos matrices columna de alturas respectivas m y n, y A una matriz (m, n), denotada por $[a_{ij}]$. El producto

$${}^tXAY = [x^1\ x^2\ \ldots\ x^m] \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}$$

incluye una sola fila y una sola columna; es un elemento de K, que queda así asociado a toda elección de las matrices X e Y para A dada. Decimos que esta es una forma bilineal, porque

$${}^tXAY = \sum_{\alpha} \sum_{\beta} x^{\alpha} a_{\alpha\beta} y^{\beta}$$

es lineal tanto con relación a las x como respecto a las y.

4. Forma cuadrática.—Cuando la matriz A es cuadrada, y si se supone $X = Y$, el producto ${}^tXAX$ se convierte en un polinomio homogéneo de segundo grado con relación a las x^i; es decir, es una forma cuadrática φ:

$$\varphi = {}^tXAX = \sum_{\alpha}\sum_{\beta} a_{\alpha\beta}x^{\alpha}x^{\beta}.$$

Demostremos que en la expresión precedente interviene solo la parte simétrica S de A. El producto es una matriz de un elemento, igual a su traspuesta; por consiguiente,

$$\varphi = {}^t({}^tXAX) = {}^tX{}^tAX.$$

De donde se deduce:

$$\varphi = {}^tX\,\frac{A + {}^tA}{2}\,X = {}^tXSX,$$

designando por S la parte simétrica de A.

Se observará la analogía de esta escritura con la notación del álgebra corriente relativa al caso de una sola variable: ax^2 puede escribirse xax, porque ${}^tx = x$.

EJERCICIOS:

1. Escríbanse en forma matricial las relaciones siguientes:

$$\begin{array}{ll} z_1 = y_1 - 2y_2, & \\ z_2 = -y_1 + 3y_2, & y_1 = x_1 - 2x_2 - 3x_3, \\ z_3 = 2y_1 + 3y_2, & y_2 = 2x_1 - x_2 + 3x_3. \end{array}$$

Dedúzcanse las expresiones de z_1, z_2, z_3 en función de x_1, x_2, x_3.

2. Hállese la matriz simétrica asociada a cada una de las formas cuadráticas

$$x^2 - 5y^2 + 3z^2 + xy - yz + 7zx \qquad \text{(cuerpo Q)},$$
$$\sqrt{2}x^2 - 3xy + (1 + \sqrt{2})y^2 \qquad \text{(cuerpo R)},$$
$$3x^2 - 2y^2 + 1z^2 - 3xy + 2yz - 1zx \qquad \text{(cuerpo } Z/5\text{)}.$$

15. Matrices elementales.—Hemos visto (pág. 140) la importancia que adquieren en el cálculo numérico ciertas combinaciones efectuadas sobre un conjunto dado de vectores. Las matrices que realizan estas operaciones sobre las filas o las columnas de una matriz A se llaman *matrices elementales*.

Razonemos sobre las filas; las matrices elementales estudiadas actúan mediante multiplicación a la izquierda, y se sabe (véase pág. 308)

que se obtienen estas matrices realizando la combinación estudiada sobre las filas de la matriz unidad cuyo orden es igual al número m de filas de A.

1.° **Matrices de permutación.**—Cambiando las filas de órdenes i y j de I_m se obtiene una matriz que, mediante multiplicación a la izquierda, realiza la permutación de las filas de órdenes i y j de A.

2.° La matriz $\mathcal{H}_\lambda = \text{diag.}\ (1, 1, \ldots, \lambda, \ldots, 1)$, deducida de la matriz unidad multiplicando una de sus filas por $\lambda(\neq 0)$ efectúa la misma operación sobre la fila correspondiente de A.

3.° Finalmente, la matriz $\mathcal{K}_\lambda = I_m + \lambda E_{ij}$, donde E_{ij} es una matriz de la base canónica, se deduce de I_m sumando a la fila i-ésima la fila j-ésima multiplicada por λ. Realiza sobre A la misma operación elemental.

4.° Con independencia de los ejemplos anteriores se pueden formar gran número de combinaciones interesantes. Señalemos, p. ej., las *matrices de sustitución*, que se deducen de I_m efectuando una sustitución sobre sus filas. El número de ellas es $m!$

EJERCICIOS:

1. Sea una matriz cuadrada J de orden n en la que todos los elementos son nulos, a excepción de los situados en la diagonal perpendicular a la diagonal principal (diagonal secundaria), que son iguales a 1. Descríbase el efecto sobre una matriz (m, n), A, de una multiplicación a la izquierda por J. Examínese igualmente el caso de una multiplicación a la derecha sobre una matriz conveniente. Calcúlense $J^2, J^3, \ldots, J^r$.

2. Descríbase el efecto, sobre una matriz A, de una multiplicación a la izquierda o a la derecha por una matriz E_j^i de la base canónica. Caso particular, $i=j$.

3. Sea A una matriz cuadrada de orden 4. Precísese cómo se deducen de A las matricees AJ y JA si es

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

16. Operaciones con cajas.—Hemos dicho que subdividir una matriz en cajas equivale a considerarla como una matriz cuyos términos son matrices. Se encuentra así el punto de vista inicial en el que los elementos pertenecían a un conjunto $\mathfrak{D}$. Para un caso tan general, no se han definido la adición y la multiplicación de matrices. Sin embargo, adoptando las mismas definiciones que para las matrices sobre un cuerpo, comprobamos que son posibles la adición y la multiplica-

ción al nivel de matrices lo mismo que lo eran al nivel de elementos[1]. Así, p. ej., supongamos que se trata de sumar dos matrices A y B descompuestas en cajas:

$$A = \begin{bmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1^1 & B_1^2 \\ B_2^1 & B_2^2 \end{bmatrix}.$$

Las restricciones acerca de las dimensiones de dos matrices sumables exigen que las cajas que ocupen lugares homólogos tengan las mismas dimensiones, y a su vez que las dos matrices A y B tengan las mismas dimensiones en sus descomposiciones en cajas. Resulta entonces de manera evidente

$$A + B = \begin{bmatrix} A_1^1 + B_1^1 & A_1^2 + B_1^2 \\ A_2^1 + B_2^1 & A_2^2 + B_2^2 \end{bmatrix},$$

El siguiente teorema permite generalizar el método al caso de la multiplicación.

TEOREMA.—*El producto por cajas de dos matrices proporciona, siempre que sea posible, una subdivisión en cajas del producto de estas matrices.*

Sea multiplicar las dos matrices

$$A = \begin{bmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

donde B_1 y B_2 son ahora cajas de filas (α y β filas, respectivamente, si A_1^1 y A_1^2 tienen α y β columnas). Consideremos la matriz

$$\begin{bmatrix} A_1^1 B_1 + A_1^2 B_2 \\ A_2^1 B_1 + A_2^2 B_2 \end{bmatrix}.$$

Es fácil volver a hallar una descomposición en cajas del producto AB. En la primera fila (de cajas), el término $A_1^1 B_1$ aporta para las γ primeras filas de AB (γ = número de filas de A_1^1) las combinaciones lineales de las α primeras filas de B; el término $A_1^2 B_2$ aporta, para las mismas filas de AB, las combinaciones que provienen de las β últimas filas de B; la segunda fila de cajas completa de manera análoga la matriz AB.

[1] El ejercicio 2 (pág. 298) y los números 14 y 15 (pág. 312), emplean esta posibilidad con varios 0 y 1 convencionales.

Finalmente, si B está también subdividida según sus columnas:

$$B = \begin{bmatrix} B_1^1 & B_1^2 & B_1^3 \\ B_2^1 & B_2^2 & B_2^3 \end{bmatrix},$$

la multiplicación a la izquierda por A siguiendo la regla ordinaria respeta la descomposición respecto a las columnas y realiza las combinaciones buscadas sobre las filas enteras de B. El lector lo comprobará inmediatamente escribiendo dicho producto.

EJERCICIOS:

1. Sean las matrices

$$A = \begin{bmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \beta & 1 \\ 0 & 0 & 1 & \beta \end{bmatrix} \quad y \quad B = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 1 & \alpha & 0 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & 1 & \beta \end{bmatrix};$$

calcúlense A^2, ABA y B^3 por cajas.

2. Una matriz cuadrada se llama *nilpotente* si es nula alguna de sus potencias. Demuéstrese que una matriz «diagonal por cajas»,

$$\left[\begin{array}{c:c:c} A & O & O \\ \hdashline O & B & O \\ \hdashline O & O & C \end{array}\right]$$

es nilpotente si, y solo si, lo son las cajas A, B y C.

3. Si $A = \begin{bmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{bmatrix}$, demuéstrese que ${}^tA = \begin{bmatrix} {}^tA_1^1 & {}^tA_2^1 \\ {}^tA_1^2 & {}^tA_2^2 \end{bmatrix}$.

ANILLO DE LAS MATRICES CUADRADAS DE ORDEN n

17. Estructura de anillo.—Hemos visto que las matrices (n, n), o matrices cuadradas de orden n, forman un conjunto cerrado para las operaciones binarias, adición y multiplicación. La primera dota al conjunto de estructura de grupo (aditivo). La segunda, asociativa y distributiva con relación a la adición, completa la estructura de anillo. De ello resulta el siguiente teorema.

TEOREMA.—*Las matrices cuadradas de orden n sobre un cuerpo* K *forman un anillo.*

No conmutatividad.— Ya se ha indicado (véase pág. 307) que, en general, es $AB \neq BA$; el anillo no es conmutativo.

Si dos matrices son tales que $AB = BA$, se dice que conmutan o que son permutables.

Si dos matrices son tales que $AB = -BA$ se dice que tienen la propiedad anticonmutativa.

Elemento unidad.— Ya se ha indicado (véase pág. 309) la presencia de un elemento neutro a la derecha y a la izquierda para la multiplicación; es la matriz unidad de orden n tal que $I_n = [\delta_j^i]$, donde δ_j^i es el símbolo de Kronecker definido en la página 137.

Divisores de cero.—No se trata de un anillo de integridad, porque la relación $AB = O$ no implica que uno al menos de los factores sea nulo. Así, p. ej., tenemos:

$$\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

EJERCICIOS:

1. Sean las matrices

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad y \quad C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Compárense los productos BC y CB entre sí y con A; igualmente para AC y CA, AB y BA.

2. Dadas las matrices

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad y \quad B = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$$

calcúlese $A^2 - 2AB - B^2$ y $(A + B - I)(A + B + I)$.

3. Demuéstrese que la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{verifica la relación } A^n = 3^{n-1}A.$$

4. Si

$$A = \begin{bmatrix} \operatorname{ch} a & \operatorname{sh} a \\ \operatorname{sh} a & \operatorname{ch} a \end{bmatrix},$$

demuéstrese que

$$A^n = \begin{bmatrix} \operatorname{ch} na & \operatorname{sh} na \\ \operatorname{sh} na & \operatorname{ch} na \end{bmatrix}.$$

5. En el álgebra de las matrices cuadradas, desarróllense $(A+B)^2$ y $(A+B)^3$. Examínese el caso en que A y B conmuten. El mismo problema para las expresiones

$$(A-B)^2(A+B) \quad y \quad (A+B)(A-B)^2.$$

6. Sea

$$A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Demuéstrese que

$$A^n=\begin{bmatrix} 1 & n & \dfrac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}.$$

7. Demuéstrese que, si las matrices A y B conmutan, lo mismo sucede con A^n y B^p, cualesquiera que sean los enteros n y p.

8. Sea la matriz $P=\lambda I+\mu D$, donde D es una matriz tal que todos sus elementos son iguales a 1, salvo los de la diagonal principal que son nulos. Determínense λ y μ de manera que $P^2=I$. Hállense todas las matrices P de orden 3 ó 4.

9. Se considera de nuevo la matriz D del problema anterior, suponiéndola de orden 3. Demuéstrese que D^n puede expresarse en la forma $D^n=\lambda_n I+\mu_n D$ y determínense las relaciones de recurrencia entre λ_n, μ_n y λ_{n-1}, μ_{n-1}. Dedúzcanse a partir de estas relaciones dos combinaciones lineales de λ_{n-1} y μ_{n-1} que se reproduzcan, multiplicadas por un factor constante, cuando se pasa de $n-1$ a n. Calcúlense λ_n y μ_n.

10. Sobre R, determínense todas las matrices

$$X=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

tales que $X^2=O$.

Resp.:

$$X=\begin{bmatrix} a & b \\ c & -b \end{bmatrix}, \quad \text{con} \quad bc+a^2=0.$$

11. Sobre R, determínense todas las matrices

$$X=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

tales que

$$X^2=\begin{bmatrix} 7 & -5 \\ -15 & 12 \end{bmatrix}.$$

Resp.:

$$\begin{bmatrix} 2 & -1 \\ -3 & 3 \end{bmatrix}, \quad \begin{bmatrix} -2 & 1 \\ 3 & -3 \end{bmatrix}, \quad \pm\frac{1}{\sqrt{13}}\begin{bmatrix} 4 & -5 \\ -15 & 9 \end{bmatrix}.$$

18. Subanillos.—Señalemos algunos subanillos importantes del anillo de las matrices de orden n.

1. Matrices diagonales.— Ya hemos indicado (pág. 297) que el conjunto de las matrices diagonales de orden n es cerrado para la adición. Se demuestra igualmente que tal conjunto es también cerrado respecto de la multiplicación:

$$\text{diag}\,(a_1, a_2, \ldots, a_n) \times \text{diag}\,(b_1, b_2, \ldots, b_n) = \text{diag}\,(a_1b_1, a_2b_2, \ldots, a_nb_n).$$

Resulta, pues, un subanillo conmutativo y con elemento unidad; pero todavía con divisores de cero.

2. Matrices escalares.—Si en una matriz diagonal son iguales todos sus elementos, se obtiene una *matriz escalar:*

$$\text{diag}\,(a,\ a,\ \ldots,\ a) = a\mathrm{I}.$$

Se deduce de la matriz unidad multiplicándola por el escalar a. El conjunto de estas matrices constituye un cuerpo isomorfo al cuerpo K donde las matrices toman sus elementos.

3. Subanillo de polinomios con relación a una matriz.—Es posible elevar una matriz cuadrada A a cualquier potencia y, por consiguiente, definir un polinomio con relación a esta matriz,

$$\mathrm{P(A)} = a_0\mathrm{I} + a_1\mathrm{A} + a_2\mathrm{A}^2 + \ldots + a_p\mathrm{A}^p,$$

eligiendo los coeficientes en K. Sea P(A) el conjunto de las matrices obtenidas de esta forma cuando se considera el conjunto de los polinomios formales. Según las reglas de cálculo para polinomios, este conjunto es cerrado respecto de la adición y de la multiplicación de matrices. Constituye un subanillo conmutativo asociado a la matriz A.

OBSERVACIÓN.—Si un polinomio $P(x)$ se anula para $x = a \in K$, el polinomio matricial correspondiente P(A) admite la raíz $\mathrm{A} = a\mathrm{I}$ y es divisible por $\mathrm{A} - a\mathrm{I}$, pero la descomposición en factores obtenida no simplifica la resolución de la ecuación matricial $\mathrm{P(A)} = \mathrm{O}$, por no tratarse de un anillo de integridad. Además, las raíces de $\mathrm{P(A)} = \mathrm{O}$ no son, en general, de esta forma particular. Se encontrarán indicaciones sobre la resolución de una ecuación matricial en la página 421, ejercicios 12, 13 y 14.

EJERCICIOS:

1. Demuéstrese que las matrices triangulares de orden n (véase pág. 297) forman un subanillo del anillo de las matrices cuadradas.

2. Demuéstrese que la matriz

$$A=\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

verifica una ecuación de segundo grado,

$$A^2+\alpha A+\beta I=0.$$

Determínense α y β.

3. Demuéstrese que toda matriz cuadrada de orden n satisface una ecuación de la forma

$$A^m+\alpha_1 A^{m-1}+\ldots+\alpha_{m-1}A+\alpha_m I=0,$$

con $m\leqslant n^2$.

4. Dada la matriz

$$A=\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

pruébese que todas las matrices B que conmutan con A son de la forma $\lambda A+\mu I$. Generalícese.

5. Sea A un elemento del anillo de las matrices de orden n. Demuéstrese que el conjunto de las matrices cuadradas del mismo orden que conmutan con A constituye un subanillo.

6. Sea A una matriz cuadrada dada, sobre un cuerpo K. Demuéstrese que el conjunto de los polinomios $P\in K[x]$ tales que $P(A)=O$ constituye un ideal de polinomios. Dedúzcase del ejercicio 3 que este ideal no es degenerado. Determínese el polinomio de grado mínimo cuando $A=O$ o cuando $A=I$.

7. Sea la matriz

$$A=\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Pruébese que las matrices $A^0=I_3$, A, A^2, ..., A^n, ... pertenecen todas a un subespacio vectorial del espacio de las matrices cuadradas de orden 3. ¿Cuál es su dimensión? Dése una base.

8. Las mismas cuestiones para la matriz

$$A=\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

19. Inversión de una matriz cuadrada.—En el anillo de las matrices cuadradas de orden n, toda matriz G o toda matriz D que

verifiquen, respectivamente,

$$GA = I, \qquad AD = I$$

(I, matriz unidad del anillo) se llaman inversas de A, a la izquierda o a la derecha. Vamos a demostrar que la *regularidad* definida en la página 301 asegura la existencia de G y de D, y que $G = D$.

Inversa a la izquierda.—Descompongamos A según sus filas A_1, A_2, ..., A_n y busquemos $G = [g_i^j]$. Se tiene:

$$\begin{bmatrix} g_1^1 & g_1^2 & \cdots & g_1^n \\ \vdots & \vdots & & \vdots \\ g_n^1 & g_n^2 & \cdots & g_n^n \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} \Sigma g_1^\alpha A_\alpha \\ \vdots \\ \Sigma g_n^\alpha A_\alpha \end{bmatrix}.$$

Si deseamos que sea $GA = I$, se debe poder determinar G de manera que los n vectores fila de la matriz del segundo miembro constituyan los vectores $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de la base canónica de K^n:

$$\mathbf{e}_i = \sum_1^n g_i^\alpha \mathbf{A}_\alpha \qquad (i \text{ de } 1 \text{ a } n).$$

Para ello es necesario que los vectores $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ engendren K^n y, puesto que su número es n, deben ser independientes y formar una base. Estas condiciones son evidentemente suficientes: si $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ constituyen una base, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ se podrá descomponer según esta base y determinar las filas sucesivas de la matriz G.

TEOREMA 1.—*Sea* A *una matriz cuadrada de orden n, sobre un cuerpo* K. *Admite una inversa a la izquierda si, y solo si, sus n vectores fila constituyen una base del espacio* K^n. *Las filas de la inversa están formadas por las coordenadas de los vectores de la base canónica referidos a la base precedente.*

Inversa a la derecha.—El lector traspondrá sin dificultad el razonamiento anterior para estudiar la inversa a la derecha. Se utiliza el segundo teorema (véase pág. 308). Enunciemos el resultado.

TEOREMA 2.—*Una matriz cuadrada de orden n admite una inversa a la derecha si, y solo si, sus n vectores columna constituyen una base del espacio* K^n. *Las columnas de la inversa están constituidas por las coordenadas de los vectores de la base canónica referidas a la base anterior.*

Equivalencia de las condiciones precedentes[1].— Supongamos que existe una inversa a la izquierda y demostremos la independencia de las columnas de A. Si fueran dependientes, al menos una de sus combinaciones lineales sería nula, es decir, existiría una matriz columna $X \neq O$, tal que

$$AX = O.$$

Multiplicando a la izquierda por G, inversa a la izquierda, se tendría:

$$G(AX) = (GA)X = IX = X,$$

de donde $X = O$, lo que contradice la hipótesis de dependencia. Las columnas son, pues, independientes y existe una inversa a la derecha, D.

Igualdad $G = D$.— Es consecuencia de la asociatividad. Tenemos

$$G = G(AD) = (GA)D = ID = D.$$

En resumen: la regularidad de una matriz asegura la existencia de su matriz inversa a la derecha y a la izquierda. Denotándola por A^{-1}, se tiene:

$$A^{-1}A = AA^{-1} = I.$$

En lo que sigue emplearemos indistintamente las locuciones *matriz regular* y *matriz invertible*.

Grupo de las matrices regulares.—Demostremos en primer lugar que estas matrices constituyen un subconjunto cerrado respecto de la multiplicación. Observemos, en efecto, que la inversa del producto AB es $B^{-1}A^{-1}$, puesto que

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I.$$

El producto es, pues, regular. Por último, la inversa de A^{-1} es A, e I es su propia inversa.

TEOREMA.—*Las matrices cuadradas regulares de orden n forman un grupo para la multiplicación matricial.*

Notemos, finalmente, que las matrices invertibles constituyen los divisores de la unidad del anillo de las matrices cuadradas.

20. Cambio de base en un espacio vectorial.—La teoría preceden-

[1] La teoría del rango demuestra la equivalencia de las condiciones sobre las filas o sobre las columnas. Sin embargo, no se utiliza aquí.

te está estrechamente ligada al estudio del cambio de base en un espacio vectorial (pág. 136). Sea $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ una base del espacio E. Hacemos corresponder al vector $\mathbf{v} = \Sigma x^a \mathbf{e}_a$ una matriz X obtenida escribiendo en una columna sus coordenadas:

$$X = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} = [x^i].$$

La escritura en columna de las coordenadas de un vector es un convenio adoptado con frecuencia y utilizado sistemáticamente en esta obra. Nos conduce a elegir en $[x^i]$ el superíndice como índice de fila (contrariamente a las notaciones anteriores), lo cual carece de importancia en el cálculo matricial.

Por el contrario, la elección de un superíndice para designar las coordenadas de un vector presenta cierto interés, sobre todo en las cuestiones donde interviene la dualidad (véase pág. 359).

Sean $\mathbf{f}_i = \Sigma a_i^a \mathbf{e}_a$ los vectores de la nueva base. A los n^2 coeficientes a_i^a hacemos corresponder una matriz P, llamada *matriz de cambio*. Las sucesivas columnas de P se forman con las coordenadas de $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ en la antigua base; el término de la fila i y columna j de P es, por consiguiente, a_j^i.

Si $\mathbf{v} = \Sigma x^i \mathbf{e}_i = \Sigma y^j \mathbf{f}_j$, las fórmulas del cambio de base

$$x^i = \Sigma a_\beta^i y^\beta \qquad \text{(véase pág. 136)}$$

pueden resumirse en una igualdad matricial única:

$$X = PY, \qquad [1]$$

donde $Y = [y^i]$ designa la nueva matriz columna de v. Una multiplicación a la izquierda por P hace pasar de la *nueva* a la *antigua* matriz de v.

Siendo las columnas de P independientes por hipótesis, una matriz de cambio es invertible, y recíprocamente. Según el teorema 2 anterior, la inversa P^{-1} es la matriz de cambio correspondiente al retorno a la antigua base.

Se llega a la misma conclusión escribiendo la igualdad [1] en la forma

$$Y = P^{-1}X.$$

Los elementos de P^{-1} son, pues, los n^2 números b_i^a definidos por

las fórmulas [3] (pág. 136) (i, índice de columna; α, de fila). Las relaciones [5] y [5 *bis*] de la página 137 expresan la doble igualdad matricial

$$PP^{-1}=P^{-1}P=I.$$

EJERCICIOS:

1. Sobre el cuerpo de los números reales, demuéstrese que la matriz $A=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ es regular. ¿Cuál es su matriz inversa? Pruébese que dos matrices A son permutables. El conjunto de estas matrices constituye un cuerpo tomando como leyes de composición la adición y la multiplicación de matrices. Pruébese que es isomorfo al cuerpo de los números complejos. ¿Qué matriz corresponde al número i?

2. Si A y B designan dos matrices cuadradas de orden n y si A es regular, la relación $AB=O$ implica $B=O$.

3. Demuéstrese que la matriz A que verifica la relación

$$A^2+A+I=O$$

admite inversa. Exprésese esta con ayuda de A.

4. Si ${}^tA$ es la traspuesta de A, demuéstrese que $({}^tA)^{-1}={}^t(A^{-1})$.

5. Si el producto de dos matrices A y B es regular, las dos matrices A y B son regulares.

6. Si $AB=BA$, se tiene que $A^{-1}B=BA^{-1}$ suponiendo A regular.

7. Demuéstrese que es condición necesaria y suficiente para que una matriz A sea invertible que el polinomio mínimo del ideal asociado a A (véase ejercicio 6, pág. 321) tenga un término constante diferente de cero.

8. Sea A la matriz tal que $A^k=O$. Demuéstrese que

$$(I-A)^{-1}=I+A+A^2+\ldots+A^{k-1}$$

9. Si A y B son dos matrices cuadradas, de órdenes n y p, cuyas inversas se designan por A^{-1} y B^{-1}, hállense las inversas de las matrices

$$\left[\begin{array}{c:c} A & O \\ \hdashline O & B \end{array}\right], \qquad \left[\begin{array}{c:c} A & C \\ \hdashline O & B \end{array}\right], \qquad \left[\begin{array}{c:c} A & O \\ \hdashline C & B \end{array}\right],$$

donde C representa una matriz de dimensiones convenientes para asegurar la descomposición en cajas.

10. Determínese la inversa de la matriz D^n estudiada en la página 319, ejercicio 9.

11. Generalícese al cálculo matricial la fórmula clásica

$$ax^2+2bx=a\left(x+\frac{b}{a}\right)^2-\frac{b^2}{a}$$

Se partirá de la expresión

$$\Phi(X) = {}^tXAX + 2{}^tBX,$$

donde B es una matriz columna, y se buscará C tal que

$$\Phi = {}^t(X+C)\,A\,(X+C) + \text{expresión independiente de } X.$$

¿Qué condición sustituye a la $a \neq 0$?

21. Cálculo de la inversa.—Sea $A = [a_j^i]$ una matriz regular de orden n. Se desea determinar las matrices fila $D_1, D_2, \ldots, D_n$ de la inversa $A^{-1} = [D_i]$. La relación $A[D_i] = I$ se traduce por las n igualdades (teorema 1, pág. 307):

$$\Sigma a_i^\alpha D_\alpha = e_i \qquad [1]$$

donde los segundos miembros designan las matrices filas de I. Se obtiene un sistema lineal cuyas incógnitas, $D_1, \ldots, D_n$, son elementos de un espacio vectorial. Los métodos clásicos de resolución, por sustitución o combinación lineal, se aplican evidentemente a un tal sistema que admite solución única cuando la matriz de los coeficientes, en este caso A, es regular. Se puede, pues, a partir del sistema [1] obtener las expresiones de $D_1, D_2, \ldots, D_n$ como combinaciones lineales de $e_1, e_2, \ldots, e_n$. Los ejemplos numéricos se resuelven fácilmente por este método. Con frecuencia la resolución se simplifica por los caracteres de simetría o de regularidad que presente A. Vamos a comprobar sobre un ejemplo la eficacia del método.

EJEMPLO.—Se desea calcular la matriz A^{-1}, inversa de

$$A = \begin{bmatrix} 3 & -2 & 0 & -1 \\ 0 & 2 & 2 & 1 \\ 1 & -2 & -3 & -2 \\ 0 & 1 & 2 & 1 \end{bmatrix}. \qquad \text{Ponemos} \qquad A^{-1} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix}$$

y formamos el sistema

$$\left.\begin{array}{r} 3D_1 - 2D_2 - D_4 = e_1, \\ 2D_2 + 2D_3 + D_4 = e_2, \\ D_1 - 2D_2 - 3D_3 - 2D_4 = e_3, \\ D_2 + 2D_3 + D_4 = e_4. \end{array}\right\}$$

Se deduce fácilmente que

$$3D_1 + 2D_3 = e_1 + e_2, \qquad D_1 + D_3 = e_3 + 2e_4,$$

de donde,

$$D_1 = e_1 + e_2 - 2e_3 - 4e_4 = [\ 1 \quad 1 \quad -2 \quad -4],$$
$$D_3 = -e_1 - e_2 + 3e_3 + 6e_4 = [-1 \quad -1 \quad 3 \quad 6].$$

Queda así un sistema con dos incógnitas:

$$\begin{cases} D_2 + D_4 = 2e_1 + 2e_2 - 6e_3 - 11e_4, \\ 2D_2 + D_4 = 2e_1 + 3e_2 - 6e_3 - 12e_4, \end{cases}$$

de donde,

$$D_2 = e_2 - e_4 = [0 \quad 1 \quad 0 \quad -1],$$
$$D_4 = 2e_1 + e_2 - 6e_3 - 10e_4 = [2 \quad 1 \quad -6 \quad -10].$$

La inversa es, por tanto,

$$A^{-1} = \begin{bmatrix} 1 & 1 & -2 & -4 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & 3 & 6 \\ 2 & 1 & -6 & -10 \end{bmatrix}.$$

OBSERVACIÓN.—El método precedente puede presentarse de manera diferente. Considerando A como matriz de una sustitución lineal, se tiene:

$$AX = Y \qquad [1]$$

donde $Y = [y_i]$ es la matriz columna asociada a $X = [x_i]$ (véase pág. 312). Si A es regular, existe una sustitución inversa que hace pasar de Y a X. En efecto, multiplicando a la izquierda por A^{-1}, resulta:

$$X = A^{-1}Y.$$

La matriz A^{-1} de esta sustitución se obtiene resolviendo las ecuaciones [1] con relación a los elementos $x_1, x_2, \ldots, x_n$. Se deducen así n formas lineales en $y_1, y_2, \ldots, y_n$ cuya matriz de los coeficientes es A^{-1}. Los cálculos son idénticos a los precedentes.

EJERCICIOS:

1. Inviértase la matriz $A = \begin{bmatrix} 1 & -a & 0 & 0 \\ 0 & 1 & -a & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Generalícese.

2. Sean $A=\begin{bmatrix} a_1 & a_2 & a_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ y $P=\begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & a_1 \\ -1 & a_1 & a_2 \end{bmatrix}$.

Demuéstrese que ${}^tA=PAP^{-1}$. ¿Qué cabe decir de PA? Generalícese.

3. Descompóngase cada una de las matrices siguientes en producto de matrices elementales:

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, \quad \begin{bmatrix} -2 & 4 \\ 5 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & 1 \\ 3 & 1 & 0 \end{bmatrix}.$$

4. Hállense las inversas de las matrices siguientes:

$$A=\begin{bmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 1 & 1 & 3 & 4 \\ 2 & -1 & 2 & 3 \end{bmatrix}; \quad B=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

¿Qué relación lineal existe entre B y B^{-1}?

5. Obténganse las inversas de las matrices siguientes, supuestas de orden n:

$$A=\begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}, \quad B=\begin{bmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \vdots & \vdots & \vdots & & \vdots \\ b & b & b & \dots & a \end{bmatrix}.$$

Exprésense A^{-1} y B^{-1} con ayuda de A y de I.

Resp.: $A^{-1}=\dfrac{1}{n-1}[A+(2-n)I]$;

$$B^{-1}=\frac{a+(n-2)b}{(a-b)[a+(n-1)b]}I-\frac{b}{(a-b)[a+(n-1)b]}A.$$

6. Sea $\epsilon=\cos\dfrac{2\pi}{n}+i\,\text{sen}\,\dfrac{2\pi}{n}$; se considera la matriz

$$A=\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \epsilon & \epsilon^2 & \dots & \epsilon^{n-1} \\ 1 & \epsilon^2 & \epsilon^4 & \dots & \epsilon^{2n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \epsilon^{n-1} & \epsilon^{2n-2} & \dots & \epsilon^{(n-1)2} \end{bmatrix}$$

Demuéstrese que nA^{-1} se deduce de A cambiando ϵ por ϵ^{-1}.

7. Determínense las inversas de las matrices triangulares siguientes:

$$A=\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}, \quad B=\begin{bmatrix} 1 & 2 & 3 & \dots & n \\ 0 & 1 & 2 & \dots & n-1 \\ 0 & 0 & 1 & \dots & n-2 \\ \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

8. Pruébese que la inversa de una matriz simétrica es simétrica.

22. Aplicación al estudio de las matrices rectangulares.—Consideremos el conjunto de las matrices de dimensiones dadas (m, n) y vamos a caracterizar en este conjunto ciertas clases que posean una propiedad común.

El mismo subespacio fila.—La condición para que dos matrices (m, n), A y B, tengan el mismo subespacio fila es una relación de equivalencia; una clase queda caracterizada por la siguiente propiedad.

TEOREMA.—*Para que dos matrices* A *y* B *tengan el mismo subespacio fila es necesario y suficiente que exista una matriz regular* P *tal que*

$$A = PB.$$

La condición es, evidentemente, suficiente porque, asociada con $B = P^{-1}A$, demuestra que cualquier fila de una de las matrices A y B es combinación lineal de las filas de la otra. Engendran, pues, el mismo subespacio.

La condición es también necesaria porque, si A y B tienen el mismo subespacio fila, este se puede referir a una base que comprenda r vectores, de donde una matriz con r filas, C_1, tal que

$$A = P'C_1$$

llamando P' a una matriz (m, r) de rango necesariamente igual a r (rango de A); sus columnas son, por tanto, independientes y las podemos completar con $m - r$ columnas suplementarias de manera que formen una matriz regular P_1 de orden m (teorema de la pág. 139). Se completa también C_1 con $m - r$ filas de ceros, de donde resulta una matriz (m, n), C. Se tiene así que $A = P_1C$.

Partiendo de B y utilizando la misma base se llega a $B = P_2C$, con P_2 regular. Puesto que $C = (P_2)^{-1}B$, se obtiene:

$$A = P_1(P_2)^{-1}B = PB$$

y P es una matriz regular. Todas las matrices de una equivalencia de este tipo tienen, en particular, el mismo rango.

Resulta un teorema análogo para las columnas cuando A se multiplica a la derecha por una matriz cuadrada regular de orden n.

El mismo rango.—La condición para que dos matrices de las mismas dimensiones tengan el mismo rango r es una relación de equi-

valencia. Caractericemos cada clase por una matriz sencilla; se puede tomar la

$$J_r = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$$

obtenida ampliando I_r, matriz unidad de orden r, con filas y columnas de ceros hasta formar una matriz (m, n). Decimos que J_r es la matriz (m, n) reducida de rango r. Precisemos cómo se deduce de ella una matriz cualquiera de la misma clase.

TEOREMA.—*Una condición necesaria y suficiente para que una matriz* (m, n), A, *sea de rango* r *es que existan dos matrices cuadradas regulares,* P *y* Q, *de órdenes respectivos* m *y* n, *tales que*

$$A = PJ_rQ,$$

donde J_r *es la matriz* (m, n) *reducida de rango* r.

Según el estudio anterior, la condición es evidentemente suficiente. Demostremos que es necesaria. Hemos visto que A puede escribirse como producto, $A = PC$, con P regular y C formada por r filas independientes seguidas de $m - r$ filas de ceros. Siempre es posible elegir las r filas anteriores de longitud n, como las r primeras filas de una matriz cuadrada regular Q, de orden n. Tenemos así:

$$C = J_rQ$$

y, por consiguiente,

$$A = PJ_rQ.$$

Consecuencias.—*Una condición necesaria y suficiente para que dos matrices* (m, n), A *y* B, *tengan el mismo rango es que existan dos matrices cuadradas invertibles* P *y* Q, *de órdenes respectivos* m *y* n, *tales que*

$$A = PBQ.$$

Sin duda, la condición es suficiente, y es también necesaria porque, si A y B tienen el mismo rango r, en virtud del teorema anterior se puede escribir:

$$A = P_1J_rQ_1, \qquad B = P_2J_rQ_2,$$

de donde

$$J_r = (P_2)^{-1}B(Q_2)^{-1}$$

y, por consiguiente,

$$A = P_1(P_2)^{-1}B(Q_2)^{-1}Q_1 - PBQ,$$

designando por P y Q dos matrices cuadradas invertibles.

EJERCICIOS:

1. Sean A y B dos matrices cuadradas de orden n y de rangos respectivos r_1 y r_2. Demuéstrese que el rango r de AB es tal que $r \geqslant r_1 + r_2 - n$.

(Los resultados anteriores demuestran que el rango de AB es igual al rango de una matriz de la forma $J_{r_1}PJ_{r_2}$, donde P es regular de orden n.)

2. Determínense todas las matrices cuadradas A de orden 3, tales que $A^2 = O$. Se demostrará que, si A no es nula, su rango debe ser igual a 1, y su traza, nula. (Utilícese, además del resultado del ejercicio anterior, la forma general de las matrices de rango 1 indicada en la pág. 300).

3. Determínense todas las matrices cuadradas A de orden 3, tales que $A^2 = I$. Se demostrará que $A \pm I$ debe ser una matriz de rango 1 y de traza ± 2 (con correspondencia de los signos).

EJERCICIOS SOBRE EL CAPITULO XIV

14-1. A las dos matrices con n columnas $A = [a_i^j]$ y $B = [b_i^j]$, asociamos la matriz $A * B$ de término general (*i-ésima* fila, *j-ésima* columna) $\sum_1^n a_i^a b_j^a$. ¿Cómo se obtiene $A * B$ a partir de A y de B? Demuéstrese que $A * B$ puede representarse como un producto ordinario. Si consideramos tres matrices, pruébese que, en general, no se verifica $(A * B) * C = A * (B * C)$.

14-2. *a)* Sea A una matriz cuadrada de orden n sobre el cuerpo de los reales. Si todos los elementos de A están acotados en valor absoluto por M, demuéstrese que los elementos de A^2 están acotados por nM^2, y los de A^k, por $n^{k-1}M^k$.

b) Se considera la matriz

$$P_k = I + \frac{A}{1} + \frac{A^2}{2!} + \ldots + \frac{A^k}{k!}$$

Pruébese que P_k tiende hacia un límite cuando k tiende a infinito (se demostrará para ello que cada elemento de P_k admite un límite). Se simboliza este límite por la notación e^A y se escribe

$$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \ldots + \frac{A^k}{k!} + \ldots$$

c) Calcúlese e^{tI}, donde t es un número real cualquiera.

d) Calcúlense e^A, e^B y e^C para las matrices

$$A=\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}, \quad B=\begin{bmatrix}0 & 0\\ -t & 0\end{bmatrix}, \quad C=\begin{bmatrix}0 & t\\ -t & 0\end{bmatrix}.$$

14-3. ¿Cuál es el límite de $A_n=\begin{bmatrix}1 & \frac{\alpha}{n}\\ -\frac{\alpha}{n} & 1\end{bmatrix}$ cuando el entero n tiende a infinito? (Hágase $\frac{\alpha}{n}=\operatorname{tg}\varphi$.) *Resp.:* $\lim A_n=\begin{bmatrix}\cos\alpha & \operatorname{sen}\alpha\\ -\operatorname{sen}\alpha & \cos\alpha\end{bmatrix}$.

14-4. La matriz $A=\begin{bmatrix}A_1 & B_1\\ A_2 & B_2\end{bmatrix}$ está descompuesta en cuatro cajas, siendo A_1 y B_2 cuadradas e invertibles. Si suponemos A igualmente invertible, ponemos $A^{-1}=\begin{bmatrix}X_1 & Y_1\\ X_2 & Y_2\end{bmatrix}$, siendo las cajas homólogas de las mismas dimensiones. Fórmense dos sistemas de ecuaciones matriciales en X_1, X_2, e Y_1, Y_2 y determínense estas matrices. En el caso particular de que las cuatro cajas A_1, B_1, A_2, B_2 sean del mismo orden e invertibles, demuéstrese que

$$A^{-1}=\begin{bmatrix}(A_1-B_1B_2^{-1}A_2)^{-1} & (A_2-B_2B_1^{-1}A_1)^{-1}\\ (B_1-A_1A_2^{-1}B_2)^{-1} & (B_2-A_2A_1^{-1}B_1)^{-1}\end{bmatrix}.$$

14-5. Consideremos dos matrices cuadradas de dos filas y dos columnas:

$$M=\begin{bmatrix}a & b\\ 0 & a\end{bmatrix}, \quad M'=\begin{bmatrix}a' & b'\\ 0 & a'\end{bmatrix},$$

donde a, b, a', b' son reales.

Fórmense las matrices producto $M\cdot M'$ y $M'\cdot M$; demuéstrese que son iguales.

2.° Se asocia al número complejo $z=a+ib$ la matriz

$$M(z)=\begin{bmatrix}a & b\\ 0 & a\end{bmatrix}.$$

Al producto $M(z)\cdot M(z')$ se asocia un número complejo u, denotado por $z\mathrm{T}z'$ Se define de esta forma sobre los números complejos una nueva multiplicación que, a los dos números complejos $z=a+ib$, $z'=a'+ib'$, asocia el producto

$$u=z\mathrm{T}z'=aa'+i(ab'+ba').$$

Demuéstrese que esta multiplicación (T) es conmutativa, asociativa y distributiva con relación a la adición ordinaria de los números complejos.

3.° Si $u=c+id$, $v=c'+id'$ son dos complejos dados, resuélvase y discútase la ecuación de primer grado

$$u\mathrm{T}z=v, \qquad [1]$$

donde z es un número complejo desconocido. Se demostrará que, si u es imaginario puro, la ecuación [1] es, en general, imposible o indeterminada; en este último caso dense todas las soluciones.

4.° Determínense todas las soluciones de la ecuación $zTz=v$; discusión. Examínese el caso particular $v=0$. Resuélvase y discútase la ecuación general de segundo grado $zTz+2uTz=v$.

5.° Se designará por $z^{[n]}$ el producto (T) de n números iguales a z:

$$z^{[n]}=zTzT\ldots z \qquad (n \text{ factores}).$$

Calcúlense $i^{[n]}$ e $(x+iy)^{[n]}$. Resuélvase y discútase la ecuación

$$z^{[n]}=u, \qquad [2]$$

donde $u=c+id$ es dado. Se demostrará que si c es estrictamente positivo, la ecuación [2] admite una solución con parte real positiva, que se designará por $\sqrt[[n]]{u}$ y se calculará en función de c y d.

¿Puede atribuirse sentido al símbolo $z^{[\alpha]}$ si α es real?

6.° Se escribe

$$E(z)=1+z+\frac{z^{[2]}}{2!}+\ldots+\frac{z^{[n]}}{n!}+\ldots$$

Demuéstrese que esta serie converge absolutamente para todo valor de z. ¿Para qué valores de z contiene solo un número finito de términos no nulos? Calcúlese explícitamente $E(z)$ en función de $z=x+iy$. Pruébese que

$$E(z+z')=E(z)TE(z').$$

7.° Resuélvase y discútase la ecuación $E(z)=u$; ¿qué relación existe entre el resultado y el límite, cuando n tiende a infinito, de

$$n(\sqrt[[n]]{u}-1)?$$

14-6. 1.° Sea $A=\begin{bmatrix}\cos a & -\operatorname{sen} a\\ \operatorname{sen} a & \cos a\end{bmatrix}$. Calcúlese A^p para p entero, positivo o negativo.

2.° Desarrollando la suma $(A+A^{-1})^p$ o la diferencia $(A-A^{-1})^p$, dedúzcanse las expresiones de $\cos^p a$ y de $\operatorname{sen}^p a$, en sumas de cosenos o senos de arcos múltiplos de a.

14-7. 1.° Siendo S una matriz de sustitución, calcúlese S^{-1}. Razónese primero sobre una matriz de orden 2 ó 3.

2.° Sea D una matriz diagonal de orden n sobre el cuerpo de los reales. Calcúlese $S^{-1}DS$, donde S designa una matriz de sustitución.

3.° Se considera el conjunto $\mathfrak{M}$ de las matrices cuadradas reales de orden n en las que cada fila y cada columna contienen un elemento distinto de cero, siendo todos los demás nulos. Demuéstrese que toda matriz $M \in \mathfrak{M}$ se escribe de una manera, y una sola, como producto DS de una matriz diagonal de determinante no nulo y de una matriz de sustitución. Pruébese, teniendo en cuenta el punto 2.°, que toda matriz $M \in \mathfrak{M}$ puede escribirse también en la forma S_1D_1, producto de una matriz de sustitución por una matriz diagonal de determinante no nulo. Demuéstrese que toda matriz $M \in \mathfrak{M}$ es invertible y determínese su inversa. Pruébese que $\mathfrak{M}$ es un grupo respecto de la multiplicación de matrices (grupo de las matrices *monomias*).

CAPITULO XV

APLICACIONES LINEALES. ESPACIOS DUALES

APLICACIONES LINEALES

1. Definiciones generales.—Sean E, F, G, ... espacios vectoriales que admiten el mismo cuerpo de escalares, K. Según la definición de la página 7, una aplicación de E en F es una correspondencia que asocia a todo vector $\mathbf{v} \in E$, un vector bien determinado $\mathbf{w} = T\mathbf{v}$, el cual pertenece a F y se llama imagen (o transformado de v). Se precisa aquí este concepto, limitándonos a considerar las aplicaciones *compatibles* con las estructuras vectoriales de E y de F.

Aplicación lineal.—Una aplicación de E en F se dice *lineal* cuando se verifican las propiedades siguientes, válidas cualesquiera que sean v y $\mathbf{v}_1$ en E:

$$T(\mathbf{v} + \mathbf{v}_1) = T\mathbf{v} + T\mathbf{v}_1, \qquad [1]$$

$$T(\lambda\mathbf{v}) = \lambda T\mathbf{v}. \qquad [2]$$

Estas propiedades fundamentales se enuncian así:

1.ª el transformado de la suma (en E) de dos vectores es la suma (en F) de sus transformados;

2.ª el transformado del homotético en E de un vector es el homotético en F del transformado de este vector, con la misma razón de homotecia. El siguiente enunciado sintetiza las dos condiciones anteriores.

TEOREMA.—*Para que una aplicación de* E *en* F *sea lineal es necesario y suficiente que conserve las combinaciones lineales:*

$$T(\lambda\mathbf{v} + \mu\mathbf{v}_1) = \lambda T\mathbf{v} + \mu T\mathbf{v}_1. \qquad [3]$$

OBSERVACIONES.—1.ª En lugar de aplicación lineal se dice muchas veces, *transformación* u *operación* lineal; utilizaremos indistintamente estos términos.

2.ª No se hace ninguna hipótesis sobre los espacios E o F, que pueden ser de dimensiones finitas o infinitas, distintos o confundidos.

3.ª En el caso $E=F$, una aplicación lineal es un *endomorfismo* (véase pág. 39) de la estructura vectorial del espacio estudiado.

2. Ejemplos de aplicaciones lineales. I. Indiquemos en primer lugar algunos ejemplos extraídos del álgebra o del análisis.

a) Supongamos $E=F=R$ (cuerpo de los reales considerado como espacio vectorial) y tratemos de caracterizar los endomorfismos de este espacio. Si el número 1 se transforma en el $a \in R$, el número $x=x\cdot 1$ se transforma en el $y=ax$ en virtud de la ley [2]; se verifica además la ley [1]. La proporcionalidad de dos sistemas de números constituye, sin duda, el ejemplo más sencillo de transformación lineal [1] y, en muchos aspectos, su estudio facilitará el del caso general [2].

b) Las siguientes fórmulas:

$$y_1=ax_1+bx_2, \qquad y_2=a'x_1+b'x_2, \qquad y_3=a''x_1+b''x_2,$$

que hacen corresponder a un vector (x_1, x_2) de R^2 un vector (y_1, y_2, y_3) de R^3, definen igualmente una aplicación lineal. El estudio de este ejemplo (y, en general, el de las aplicaciones de R^n en R^m) se hará más adelante.

c) Sea E el espacio de las funciones derivables, con derivada continua en un intervalo dado de R, y sea F el espacio de las funciones continuas en el mismo intervalo. La aplicación que hace corresponder a f su derivada f' es lineal.

d) Sea E el espacio de las funciones continuas en el segmento [0, 1]. La aplicación $f \longrightarrow \int_0^1 f(x)dx$ de E en R es lineal.

e) Si E_1 y E_2 son dos subespacios suplementarios de E, un vector $\mathbf{v}$ cualquiera se escribe $\mathbf{v}=\mathbf{v}_1+\mathbf{v}_2$ con $\mathbf{v}_1 \in E_1$ y $\mathbf{v}_2 \in E_2$. La correspondencia $\mathbf{v} \longrightarrow \mathbf{v}_1$ es una aplicación lineal de E en E_1.

II. La geometría ordinaria proporciona también numerosos ejemplos de aplicaciones lineales. Asociamos al espacio puntual un espacio vectorial, que se obtiene al unir un mismo punto O (origen) a los demás puntos del espacio.

Sea una transformación puntual $M'=T(M)$ que conserva el punto O. Se la puede considerar como una transformación que actúa sobre el vector $\mathbf{OM}$ y lo transforma en $\mathbf{OM'}$. Tal transformación es lineal cuando conserva las combinaciones lineales:

$$T(\lambda\mathbf{OM}+\mu\mathbf{ON})=\lambda T(\mathbf{OM})+\mu T(\mathbf{ON}).$$

[1] Es conveniente observar que la relación $y=ax+b$ (función lineal de la enseñanza elemental) no define una transformación lineal.

[2] Así, p. ej., la escritura $Y=AX$ del capítulo 17 generaliza la $y=ax$.

Se comprueba fácilmente que numerosas transformaciones estudiadas en geometría elemental satisfacen esta condición; p. ej.:

a) la simetría con relación al origen: $\mathbf{OM'} = -\mathbf{OM}$;

b) la homotecia de centro O y razón k: $\mathbf{OM'} = k\mathbf{OM}$;

c) la proyección de M sobre el plano xOy paralelamente a Oz;

d) la rotación de ángulo α alrededor de un eje orientado Δ que pasa por O;

e) la semejanza directa que conserva un eje Δ que pasa por O.

3. Imagen de un sistema.—Sea $S = \{\mathbf{v}_i\}$ un sistema de vectores de E (i varía de 1 a p). En la transformación se convierte en el sistema $TS = \{T\mathbf{v}_i\}$ de F. Indiquemos algunas consecuencias inmediatas de la fórmula [3].

a) La imagen del vector nulo de E es el vector nulo de F. En efecto, la igualdad

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{O}_E$$

implica

$$T(\mathbf{O}_E) = T(\mathbf{v}) + T(-\mathbf{v}) = T\mathbf{v} - T\mathbf{v} = \mathbf{O}_F .$$

b) La imagen de un sistema ligado de E es un sistema ligado de F, ya que toda relación de dependencia entre los $\mathbf{v}_i$ se conserva en la transformación. Si S es ligado se tiene, al menos, una relación

$$\lambda^1\mathbf{v}_1 + \lambda^2\mathbf{v}_2 + \ldots + \lambda^p\mathbf{v}_p = \mathbf{O}_E,$$

que, en la transformación, se convierte en

$$\lambda^1 T\mathbf{v}_1 + \lambda^2 T\mathbf{v}_2 + \ldots + \lambda^p T\mathbf{v}_p = T\mathbf{O}_E = \mathbf{O}_F;$$

por consiguiente, el sistema TS es ligado.

c) Si la imagen TS de un sistema es libre en F, el sistema S es libre en E. La propiedad es consecuencia inmediata de la proposición anterior.

d) La imagen de un sistema libre de E puede ser un sistema ligado de F. Así, p. ej., en la transformación $y = ax_1 + bx_2$, que asocia a todo vector (x_1, x_2) de R^2 el vector $y \in R$, dos vectores independientes de R^2 tienen por imagen dos vectores dependientes de R (de dimensión igual a 1).

e) De manera general, sea P una parte cualquiera de E. Los transformados $T\mathbf{v}$ de los vectores $\mathbf{v} \in P$ constituyen una parte de F, denotada por T(P) y que se llama imagen de P.

4. Determinación de una aplicación lineal. Existencia.—La definición de aplicación lineal se ha establecido *a priori*. El empleo de una base permite demostrar la existencia de tales aplicaciones y precisar su determinación.

Existencia y unicidad. TEOREMA.—*Existe una aplicación lineal y una sola de* E *en* F *tal que los transformados de los vectores de una base de* E *sean vectores de* F *arbitrariamente elegidos.*

Sea B [1] una base de E y elijamos arbitrariamente el transformado **f** de todo $\mathbf{e} \in B$. Demostremos que de esta manera se define una transformación lineal T de E en F. Un vector $\mathbf{v} \in E$ es combinación lineal de un número finito (no conocido *a priori*) de elementos de B. Numerando estos elementos, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p$, se tiene:

$$\mathbf{v} = \sum_1^p x^\alpha \mathbf{e}_\alpha. \qquad [1]$$

La propiedad fundamental exige que el transformado de **v** sea

$$\mathbf{w} = \sum_1^p x^\alpha \mathbf{f}_\alpha,$$

el cual está bien determinado; se demuestra sin dificultad el carácter lineal de la aplicación de E en F así obtenida.

La unicidad de la aplicación que satisface a las condiciones dadas resulta, finalmente, de la unicidad de la descomposición [1] y del carácter necesario de la expresión elegida para Tv.

5. Imagen de un subespacio. Rango. Aplicación suprayectiva. TEOREMA 1.—*La imagen de un subespacio de* E *es un subespacio de* F.

Sea E_1 un subespacio de E. Demostremos que su imagen $T(E_1)$ es un subespacio de F. Si $\mathbf{w}_1$ y $\mathbf{w}_2$ pertenecen a $T(E_1)$, se puede hallar $\mathbf{v}_1$ y $\mathbf{v}_2$ en E_1 tales que

$$\mathbf{w}_1 = T\mathbf{v}_1, \qquad \mathbf{w}_2 = T\mathbf{v}_2.$$

Por consiguiente, $\mathbf{w}_1 + \mathbf{w}_2 = T\mathbf{v}_1 + T\mathbf{v}_2 = T(\mathbf{v}_1 + \mathbf{v}_2)$. Pero, $(\mathbf{v}_1 + \mathbf{v}_2) \in E_1$ y su transformado $\mathbf{w}_1 + \mathbf{w}_2$ pertenece a $T(E_1)$, que es, por consiguiente,

[1] Nos colocamos en el caso general admitiendo el teorema de la página 132; el lector podrá limitarse al caso de una base finita, o infinita numerable. Los vectores de la base se ordenan de una vez para siempre.

cerrado respecto de la suma en F. Igualmente,

$$\lambda w_1 = \lambda T v_1 = T(\lambda v_1) \in T(E_1)$$

y el teorema queda demostrado.

Caso particular.—En particular, la imagen de E completo es un subespacio de F denotado por T(E) y llamado *subespacio imagen* de la transformación. Si $T(E) = F$, la aplicación es suprayectiva. El siguiente ejemplo, $y_1 = a_1 x$, $y_2 = a_2 x$, que a $x \in R$ hace corresponder $(y_1, y_2) \in R^2$, demuestra que T(E) puede ser una parte, en sentido estricto, de F; la imagen de R completo es, en efecto, el subespacio $(a_1 x, a_2 x) = x(a_1, a_2)$, de dimensión uno, de R^2.

Rango.—La dimensión de T(E), si existe, se llama *rango* de la aplicación; tanto E como F pueden ser de dimensión finita o infinita.

TEOREMA 2.—*Si* E_1 *es el subespacio engendrado por una parte* P *de* E, *la imagen* T(P) *de la parte considerada engendra el subespacio* $T(E_1)$.

Todo $w \in T(E_1)$ es imagen de al menos un $v \in E_1$, el cual es combinación de un número finito de vectores de P. Por consiguiente, $w = Tv$ es combinación lineal de las imágenes de estos vectores.

Aplicación suprayectiva.—*Para que una aplicación lineal* T, *de* E *en* F, *sea suprayectiva es necesario y suficiente que los transformados de los vectores de una base de* E *constituyan un sistema de generadores de* F.

No hacemos ninguna hipótesis sobre las dimensiones, y sea B una base de E. El teorema anterior demuestra que T(B) engendra el subespacio imagen T(E). Si la transformación es suprayectiva, $T(E) = F$ y T(B) engendra F. Recíprocamente, si engendra F, todo vector de F será combinación lineal de un número finito de vectores de T(B) y, por consiguiente, transformado de un vector de E; la transformación es suprayectiva.

6. Imagen recíproca de un subespacio de F. Núcleo. TEOREMA.—*La imagen recíproca* $T^{-1}(F_1)$ *de un subespacio cualquiera* F_1 *de* F *es un subespacio* E_1 *de* E.

En efecto: F_1 contiene al menos el vector nulo de F, que admite al menos un antecedente, el vector nulo de E; por tanto, $T^{-1}(F_1)$ no es vacío. Sean, pues, v_1 y v_2 dos vectores cualesquiera elegidos en este subconjunto. Su suma, $v_1 + v_2$, tiene por transformado $Tv_1 - Tv_2 \in F$.

y, por consiguiente, $\mathbf{v}_1+\mathbf{v}_2 \in T^{-1}(F_1)$; igualmente, $\lambda\mathbf{v}_1 \in T^{-1}(F_1)$. Hemos obtenido así un subespacio de E.

Núcleo.—Consideremos, en particular, el subespacio de F que se reduce a solo el vector nulo $\mathbf{O}_F$. Su imagen recíproca es un subespacio importante de E denominado *núcleo* de la transformación. El núcleo está, pues, constituido por el conjunto de los vectores de E aplicados sobre el vector nulo de F. Estos vectores constituyen el conjunto de soluciones de la ecuación

$$T(\mathbf{v})=\mathbf{O}_F,$$

en la que $\mathbf{v}$ es la incógnita.

El teorema siguiente pone en evidencia la importancia de este concepto.

TEOREMA.—*Una condición necesaria y suficiente para que una transformación lineal* T *sea inyectiva es que el núcleo se reduzca exclusivamente al vector nulo de* E.

En efecto, T es inyectiva si todo elemento de T(E) no tiene más que un original. En particular, al elemento nulo $\mathbf{O}_F$ debe corresponder solo un elemento de E; es decir, que $T^{-1}(\mathbf{O}_F)$ debe reducirse a $\mathbf{O}_E$.

Recíprocamente, si se verifica esta condición, dos elementos distintos de T tienen imágenes diferentes, ya que $T(\mathbf{v})=T(\mathbf{v}')$ implicaría

$$T(\mathbf{v}-\mathbf{v}')=\mathbf{O}_F,$$

de donde $\mathbf{v}-\mathbf{v}'=\mathbf{O}_E$ y $\mathbf{v}=\mathbf{v}'$, con lo que los elementos $\mathbf{v}$ y $\mathbf{v}'$ no serían distintos.

CONSECUENCIA.—*Si el núcleo se reduce al vector nulo de* E, *la aplicación de* E *sobre* T(E) *es biyectiva. Se denomina regular.*

Según la propia definición, T(E) es, a la vez, inyectiva y suprayectiva; por consiguiente, es una biyección. Existe una aplicación recíproca, T^{-1}, de T(E) en E, de la que fácilmente se demuestra que es lineal. Así, p. ej.,

$$T^{-1}(T\mathbf{v}+T\mathbf{v}_1)=T^{-1}T(\mathbf{v}+\mathbf{v}_1)=\mathbf{v}+\mathbf{v}_1=T^{-1}(T\mathbf{v})+T^{-1}(T\mathbf{v}_1).$$

Sin embargo, debe observarse que la aplicación T^{-1} se efectúa de T(E) en E y no (en general) de F en E. Ahora bien, T(E) no es conocido *a priori* porque depende de T. Al contrario, en el estudio de la página 344, se considera una biyección entre dos espacios dados *a priori*.

TEOREMA.—*Una aplicación inyectiva de* E *en* F *transforma un sistema libre de* E *en un sistema libre de* F.

Sea $S=\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ un sistema libre de E. Si el sistema $TS=\{T\mathbf{v}_1, \ldots, T\mathbf{v}_p\}$ no fuera libre, existiría al menos una combinación lineal nula, con coeficientes no todos nulos:

$$\lambda^1 T\mathbf{v}_1 + \lambda^2 T\mathbf{v}_2 + \ldots + \lambda^p T\mathbf{v}_p = \sum \lambda^i T\mathbf{v}_i = T\left(\sum \lambda^i \mathbf{v}_i\right) = \mathbf{O}_F.$$

El vector $\sum \lambda^i \mathbf{v}_i$, cuyo transformado es nulo, pertenecería al núcleo que, en este caso, se reduce a $\mathbf{O}_E$, y la relación $\sum \lambda^i \mathbf{v}_i = \mathbf{O}_E$ exigiría que S fuera ligado, lo cual es imposible.

COROLARIO.—*Una aplicación inyectiva transforma un subespacio de* n *dimensiones de* E *en un subespacio de* n *dimensiones de* F.

Tal aplicación conserva, por tanto, la dimensión.

7. Ecuaciones lineales.—La definición dada de aplicación lineal permite generalizar la noción de *ecuación lineal*. Sea T una aplicación de E en F e $\mathbf{y}_0$ un elemento dado de F. La expresión

$$T\mathbf{x} = \mathbf{y}_0, \qquad [1]$$

en la que $\mathbf{x}$ designa un elemento desconocido de E, constituye una ecuación lineal. Resolver tal ecuación consiste en determinar el conjunto de vectores de E que, en la transformación T, se aplican sobre un vector dado de F. En general, un conjunto como el anterior no es un subespacio de E.

Supongamos dados E y F. Una ecuación lineal depende de la manera de elegir T y el segundo miembro $\mathbf{y}_0$ en F. En particular, para $\mathbf{y}_0 = \mathbf{O}_F$ la ecuación se llama *homogénea*. Se escribe:

$$T\mathbf{x} = \mathbf{O}_F \qquad [2]$$

y, por lo mismo, T, se dice asociada a la ecuación [1]. Indiquemos de manera general las propiedades de las ecuaciones lineales.

Ecuación homogénea.—La ecuación $T\mathbf{x} = \mathbf{O}_F$ se ha estudiado ya en la página 339; el conjunto de sus soluciones constituye el núcleo de la aplicación T. En este caso, y solo en este caso, el conjunto de las soluciones es un subespacio de E. Una ecuación homogénea admite siempre una solución al menos: $\mathbf{x} = \mathbf{O}_E$ (solución trivial).

Caso general.—Cualquiera que sea $\mathbf{x} \in E$, se sabe que T(x) pertenece al subespacio imagen T(E). Por consiguiente, una condición necesaria

y suficiente de posibilidad es que $y_0 \in T(E)$. Suponiendo que se verifica esta condición, la ecuación admite al menos una solución, x_0. Precisemos el conjunto de sus soluciones.

TEOREMA.—*Se obtiene el conjunto de todas las soluciones de la ecuación* [1] *añadiendo a una solución particular* x_0 *de esta ecuación una cualquiera de las soluciones de la ecuación homogénea asociada* [2]:

$$x = x_0 + T^{-1}(O_F).$$

Haciendo $x = x_0 + x'$, en virtud de la linealidad, se tiene:

$$T(x_0 + x') = Tx_0 + Tx';$$

ya que $Tx_0 = y_0$, queda $Tx' = O_F$. Se ve que x' es una solución cualquiera de la ecuación homogénea asociada [2].

Suponiendo verificada la condición de posibilidad, se deduce que la ecuación [1] tendrá solución única si, y solo si, la ecuación homogénea asociada admite solo la solución trivial.

COMPOSICION DE APLICACIONES LINEALES

8. Espacio vectorial de las aplicaciones de E en F.—Sean dos espacios E y F de un mismo cuerpo K. Sobre el conjunto de las aplicaciones lineales de E en F se pueden definir ciertas leyes de composición que dotan a este conjunto de estructura de espacio vectorial.

Suma.—Se llama suma de las aplicaciones T y T_1, la aplicación T_2 que, a todo $v \in E$, asocia el vector de F definido por la igualdad

$$T_2(v) = T(v) + T_1(v).$$

La imagen de v por T_2 es, por consiguiente, la suma, en F, de las imágenes de v por T y T_1. Se escribe:

$$T_2 = T + T_1.$$

La igualdad

$$\begin{aligned} T_2(\lambda v + \mu v') &= T(\lambda v + \mu v') + T_1(\lambda v + \mu v') \\ &= \lambda[T(v) + T_1(v)] + \mu[T(v') + T_1(v')] \\ &= \lambda T_2(v) + \mu T_2(v') \end{aligned}$$

demuestra que T_2 es una aplicación lineal.

La conmutatividad y la asociatividad de la adición de aplicaciones son consecuencias inmediatas de las propiedades análogas sobre E y F. Existe una aplicación neutra para la adición: es la aplicación nula, O, que a todo $v \in E$ asocia $O(v) = 0_F$. Finalmente, la aplicación opuesta de T es $(-T)$ tal que $(-T)(v) = -T(v)$ para todo v; se tiene que $T + (-T) = O$. Esta operación dota así de estructura de grupo aditivo al conjunto de las aplicaciones de E en F.

Multiplicación por un escalar.—Sea una aplicación T y un escalar $\alpha \in K$; les asociamos la aplicación αT, tal que para todo v se tenga

$$(\alpha T)(v) = \alpha T(v).$$

Por consiguiente, la imagen por αT es la homotética de la imagen para la aplicación T, siendo α la razón de homotecia.

La comprobación de las leyes de la estructura de espacio vectorial se reduce fácilmente a las propiedades de los espacios E y F; lo dejamos como ejercicio a la atención del lector.

TEOREMA.—*El conjunto de las aplicaciones lineales de un espacio* E, *en un espacio* F, *puede, a su vez, ser dotado de estructura de espacio vectorial.*

El estudio de este espacio se proseguirá en un caso particular en el capítulo 17.

9. Compuesta de dos aplicaciones lineales.—Consideremos tres espacios vectoriales E, F, G sobre un mismo cuerpo. Sea T una aplicación lineal de E en F, y $\mathfrak{S}$, otra aplicación lineal de F en G. Las nociones generales estudiadas en el capítulo 2 (pág. 33) nos permiten considerar la aplicación $\mathfrak{S} \circ T$ de E en G, que hace corresponder a todo vector $v \in E$ el vector $\mathfrak{S}[T(v)]$. Se llama compuesta de las aplicaciones T y $\mathfrak{S}$.

TEOREMA 1.—*La compuesta de dos aplicaciones lineales es lineal.*

En efecto,

$$\mathfrak{S}[T(\lambda v + \mu v_1)] = \mathfrak{S}[\lambda T(v) + \mu T(v_1)] = \lambda \mathfrak{S}[T(v)] + \mu \mathfrak{S} T[(v_1)].$$

TEOREMA 2.—*La composición de aplicaciones es distributiva con relación a las operaciones sobre los espacios T o* $\mathfrak{S}$.

Comprobemos, p. ej., que

$$\mathfrak{S} \circ (\lambda T + \mu T_1) = \lambda \mathfrak{S} \circ T + \mu \mathfrak{S} \circ T_1.$$

En primer lugar, para un vector $\mathbf{v}$ cualquiera, se tiene:

$(\lambda T+\mu T_1)(\mathbf{v})=\lambda T(\mathbf{v})+\mu T_1(\mathbf{v})$ (operaciones sobre el conjunto de las T),

de donde,

$$\begin{aligned}\mathfrak{G}[\lambda T+\mu T_1](\mathbf{v}) &= \mathfrak{G}[\lambda T(\mathbf{v})]+\mathfrak{G}[\mu T_1(\mathbf{v})] \quad \text{(propiedad de } \mathfrak{G}\text{)}\\ &= \lambda\mathfrak{G}[T(\mathbf{v})]+\mu\mathfrak{G}[T_1(\mathbf{v})].\end{aligned}$$

Caso de los endomorfismos.—Supongamos $E=F=G$. Se trata de las aplicaciones de E en E que conservan la estructura de espacio vectorial; es decir, de los endomorfismos de un espacio vectorial. Según el apartado anterior, forman un espacio vectorial. La composición introducida anteriormente define una multiplicación sobre el conjunto, operación que, como se sabe, es asociativa. La distributividad, ya demostrada, enseña que el conjunto estudiado tiene estructura de anillo. Se trata de un anillo no conmutativo, porque, en general,

$$\mathfrak{G}\circ T \neq T\circ \mathfrak{G}.$$

Existe un elemento unidad: la transformación idéntica I.

APLICACIONES ENTRE ESPACIOS DE DIMENSIONES FINITAS

10. El espacio inicial es de n dimensiones.—Los resultados anteriores son independientes del concepto de dimensión. Supongamos ahora que E sea de n dimensiones, y F, cualquiera.

El subespacio imagen, engendrado por los transformados de los vectores de una base de E, tiene dimensión finita (rango) inferior o igual a n; por consiguiente, una aplicación lineal puede solo disminuir (o conservar) la dimensión del espacio inicial.

Por otra parte, el núcleo $T^{-1}(\mathbf{O}_F)$ es un subespacio de E; tiene también dimensión inferior o igual a n. El siguiente teorema relaciona las dimensiones de estos dos subespacios.

TEOREMA FUNDAMENTAL.—*Toda aplicación lineal de un espacio de n dimensiones en un espacio cualquiera verifica la condición:*

$$\textit{dimensión del núcleo} + \textit{rango} = n.$$

Sean s la dimensión del núcleo y $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_s\}$ una base de este subespacio. Es posible hallar $n-s=r$ vectores $\{\mathbf{e}_{s+1}, \ldots, \mathbf{e}_n\}$ que en-

gendren un subespacio suplementario. El sistema $\{\mathbf{e}_i\}$ (i de 1 a n) constituye una base de E. Los transformados $\{\mathrm{T}\mathbf{e}_i\}$ forman un sistema de generadores de T(E). Pero, $\mathrm{T}\mathbf{e}_i = \mathbf{O}_F$ (i de 1 a s) y $\mathrm{T}\mathbf{e}_{s+1}, \ldots, \mathrm{T}\mathbf{e}_n$ engendran T(E). Demostremos que son independientes. Una igualdad

$$\lambda^1 \mathrm{T}\mathbf{e}_{s+1} + \lambda^2 \mathrm{T}\mathbf{e}_{s+2} + \ldots + \lambda^{n-s} \mathrm{T}\mathbf{e}_n = \mathbf{O}_F$$

exigiría que el vector $\lambda^1 \mathbf{e}_{s+1} + \ldots + \lambda^{n-s} \mathbf{e}_n$ perteneciese al núcleo, lo que es imposible (salvo para $\lambda^1 = \lambda^2 = \ldots = \lambda^{n-s} = 0$), puesto que pertenece a un subespacio suplementario. Se tiene así, efectivamente, una base de T(E), que es de $n - s = r$ dimensiones.

En particular, para que la aplicación considerada sea inyectiva, es necesario y suficiente que el rango sea igual a n, dimensión del espacio inicial. Por último, la hipótesis de que el espacio F posee dimensión finita, m, presenta menos interés que la anterior. El subespacio imagen tiene ciertamente dimensión finita r, y $r \leqslant m$. Para $r = m$, la transformación es suprayectiva.

11. Biyección.—Consideremos dos espacios E y F de dimensiones respectivas n y m y tratemos de caracterizar las biyecciones de E en F. La transformación debe ser inyectiva y su rango r ha de ser igual a n. Debe ser también suprayectiva, y su rango, igual a m. Por tanto, $r = n = m$.

TEOREMA 1.—*Una biyección puede establecerse solo entre dos espacios de la misma dimensión.*

Encontramos de nuevo la proposición de la página 134: para que dos espacios vectoriales sean isomorfos es necesario y suficiente que tengan la misma dimensión.

TEOREMA 2.—*Entre dos espacios de la misma dimensión toda aplicación lineal inyectiva (suprayectiva) es también suprayectiva (inyectiva). Por consiguiente, es una biyección.*

En el primer caso (inyectiva), por ser el rango igual a la dimensión común de los dos espacios, la aplicación es suprayectiva.

En el segundo (suprayectiva), el rango es igual a la dimensión común y, por consiguiente (teorema fundamental), el núcleo se reduce al vector nulo únicamente; la aplicación es inyectiva.

TEOREMA 3.—*Una condición necesaria y suficiente para que una aplicación lineal entre dos espacios de dimensiones finitas sea una biyección es que transforme una base del primero en una base del segundo.*

Los transformados de los n vectores de una base de E forman una base de F, $n = m$; la transformación es suprayectiva.

EJERCICIOS:

1. Sea T un endomorfismo de un espacio E. Demuéstrese que el conjunto de los vectores $v \in E$ que verifican la condición $Tv=v$ es un subespacio de E. Precísese la intersección de este subespacio y del núcleo de T.

2. Un endomorfismo T de E que verifica la condición $T^2=T \circ T=T$ se llama *proyección* (o transformación idempotente). ¿Qué propiedad poseen los vectores del subespacio imagen de T? Dedúzcase de ella que el núcleo de una proyección y el subespacio imagen son suplementarios. Se vuelve a encontrar así la definición de proyección dada en la página 125 (extendida, sin embargo, al caso en que $T=I$).

3. Demuéstrese que una condición necesaria y suficiente para que un endomorfismo H sea permutable con un endomorfismo cualquiera T ($T \circ H = H \circ T$) es que H sea una homotecia: $Hv=\lambda v$, con λ constante arbitraria.

Se demostrará en primer lugar que H transforma todo vector v en un vector proporcional, y después que el coeficiente de proporcionalidad es independiente del vector elegido.

4. Sea E de dimensión n y F de dimensión m; supóngase que T es una aplicación de E en F y N el núcleo de T.

a) Si E_1 es un subespacio de E de dimensión p, demuéstrese que su imagen $T(E_1)$ es un subespacio de F de dimensión $p-\dim(E_1 \cap N)$.

b) Sea $T(E)=I$ el subespacio imagen de la transformación, y r, su dimensión. Si F_2 es un subespacio de F, pruébese que su imagen recíproca E_2 en E es tal que, $\dim E_2=n-r+\dim(F_2 \cap I)$.

5. Sea $E=E_1 \oplus E_2 \oplus \ldots \oplus E_p$ una descomposición en suma directa del espacio E. Si $v=\sum_1^p v_i$ con $v_i \in E_i$, demuéstrese que la aplicación T_i de E en E que a v asocia $T_i v=v_i$ es un endomorfismo. Precísense el subespacio imagen y el núcleo de T_i.

Caracterícese la suma de los endomorfismos $\sum_1^p T_i$.

Estúdiese el producto $T_j \circ T_i$ de dos de estos endomorfismos, distintos o no.

6. Un endomorfismo T se llama recíproco o involutivo cuando su cuadrado T^2 es la transformación idéntica. Si T_1 es una proyección, pruébese que $T=2T_1-I$ es recíproca; recíprocamente, demuéstrese que toda transformación recíproca puede ponerse en esta forma, suponiendo el cuerpo de los escalares de E de característica no igual a 2.

EL ESPACIO DE LAS FUNCIONES LINEALES

12. Funciones lineales.—Sea E un espacio vectorial sobre un cuerpo K. Se sabe que K es también un espacio vectorial (de una dimensión) y podemos estudiar las aplicaciones lineales de E en K, lo que

constituye un caso particular de la teoría anterior. Comencemos por precisar las definiciones. Una aplicación f de un espacio E en su cuerpo de escalares es una correspondencia que, a todo vector $\mathbf{v} \in E$, asocia $f(\mathbf{v}) \in K$. La aplicación es lineal cuando conserva las combinaciones lineales; es decir,

$$f(\lambda \mathbf{v} + \mu \mathbf{v}') = \lambda f(\mathbf{v}) + \mu f(\mathbf{v}')$$

cualesquiera que sean los escalares λ y μ y los vectores $\mathbf{v}$ y $\mathbf{v}'$. En particular, al vector nulo corresponde siempre el escalar nulo de K. Para subrayar que la aplicación se hace en el cuerpo de los escalares, llamaremos *función lineal* definida en E a tal aplicación. Nada se ha supuesto respecto de la dimensión de E, pero en lo que sigue nos limitamos al caso de un espacio de n dimensiones. Se puede entonces aplicar el teorema de la página 343, lo que da:

$$\text{rango} + \text{dimensión del núcleo} = n.$$

Puesto que K es de dimensión 1, cabe hacer dos hipótesis sobre el rango.

a) El rango es nulo.—Todo vector de E se aplica sobre el número 0 de K. Es la aplicación nula, $f = O$, cuyo núcleo coincide con E.

b) El rango es igual a 1.—La aplicación es suprayectiva y la dimensión del núcleo es $n-1$; es decir, el conjunto de los vectores $\mathbf{v}$ tales que $f(\mathbf{v}) = 0$ constituye un subespacio de $n-1$ dimensiones de E. Resumamos estas propiedades.

TEOREMA.—*El núcleo de una función lineal, no idénticamente nula, definida en un espacio de n dimensiones* E, *es un subespacio de* $n-1$ *dimensiones de* E.

13. Expresión analítica en una base. Existencia.—La definición de las funciones lineales se ha hecho *a priori*. El empleo de una base nos va a permitir la construcción de tales funciones.

Sea $\{\mathbf{e}_i\}$ la base y $\mathbf{v} = \sum x^\alpha \mathbf{e}_\alpha$ un vector de E. Si f es una función lineal *dada*, tendremos en virtud de la propiedad fundamental,

$$f(\mathbf{v}) = f\left(\sum x^\alpha \mathbf{e}_\alpha\right) = x^1 f(\mathbf{e}_1) + x^2 f(\mathbf{e}_2) + \ldots + x^n f(\mathbf{e}_n). \qquad [1]$$

Los valores $f(\mathbf{e}_i)$ (i de 1 a n) están bien determinados si f es conocida, y $f(\mathbf{v})$ se expresa como una *forma lineal* de las coordenadas de $\mathbf{v}$.

TEOREMA.—*Una función lineal f del vector $\mathbf{v} \in \mathrm{E}$ se expresa mediante una forma lineal de las coordenadas de* $\mathbf{v}$ *en una base cualquiera. Los coeficientes de la forma son los valores de la función para los vectores de la base.*

Se deducen fácilmente las dos propiedades siguientes:

Existencia.—*Sea* $\{\mathbf{e}_i\}$ *una base del espacio E. La aplicación, que a* $\mathbf{v} = \sum x^\alpha \mathbf{e}_\alpha$ *asocia el escalar* $f(\mathbf{v}) = \sum a_\alpha x^\alpha$, *donde las* a_α *son n escalares dados de* K, *es lineal y tal que* $f(\mathbf{e}_i) = a_i$ (i *de* 1 *a* n).

La fórmula

$$f(\mathbf{v}) = \sum a_\alpha x^\alpha$$

define ahora f; la linealidad es consecuencia inmediata, y para $\mathbf{v} = \mathbf{e}_i$ se tiene que $x^\alpha = \delta_i^\alpha$ (símbolo de Kronecker), de donde:

$$f(\mathbf{e}_i) = \sum a_\alpha \delta_i^\alpha = a_i.$$

Unicidad.—*Existe una función lineal y una sola que toma n valores dados* $(a_1, a_2, \ldots, a_n)$ *para los n vectores de una base.*

El primer teorema exige que la función buscada sea de la forma $f(\mathbf{v}) = a_1 x^1 + a_2 x^2 + \ldots + a_n x^n$, y el segundo asegura que la función así elegida verifica todas las condiciones enunciadas.

Funciones y formas.—El método anterior proporciona el modo de obtener todas las funciones lineales definidas sobre E. Elegida una base, toda función se expresa mediante una forma, y recíprocamente.

TEOREMA.—*El conjunto de las funciones lineales está representado de manera biunívoca por el conjunto de las formas lineales de las coordenadas de un vector referido a una base dada de* E.

Se ha utilizado un sistema de referencia particular, pero el valor $f(\mathbf{v})$ se halla ligado *intrínsecamente* a $\mathbf{v}$. A continuación se estudian las consecuencias de un cambio de base.

14. Cambio de base.—Una función lineal invariante queda definida por sus coeficientes $(a_1, a_2, \ldots, a_n)$ en una base $\{\mathbf{e}_i\}$ de E. Hallemos los nuevos coeficientes de esta función cuando se efectúa un cambio de base cuya matriz de cambio sea P.

Utilicemos la notación matricial. En la igualdad

$$f(\mathbf{v}) = a_1 x^1 + a_2 x^2 + \ldots + a_n x^n$$

se reconoce el producto de la matriz columna

$$X=\begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}$$

del vector v por una matriz fila $A=[a_1, a_2, \ldots, a_n]$ que caracteriza a la función:

$$f(\mathbf{v})=AX.$$

Después de efectuado el cambio de base, v toma una nueva matriz X′ tal que $X=PX'$, y como $f(\mathbf{v})$ permanece invariable, se tiene:

$$f(\mathbf{v})=A(PX')=(AP)X'=A'X',$$

donde $A'=AP$. Se ve que A′ es la nueva matriz de los coeficientes de la función f.

TEOREMA.—*En un cambio de base, de matriz de cambio* P, *la matriz fila* A *de una función lineal queda multiplicada por la matriz* P.

15. Operaciones con las funciones lineales.—Hasta el momento hemos estudiado cada función lineal individualmente. Consideremos ahora, en cambio, el conjunto F de estas funciones. Para formarlo, basta elegir una base de E y atribuir n valores arbitrarios $(a_1, a_2, \ldots, a_n)$ a los n vectores de la base. El conjunto de las elecciones de estos n coeficientes nos proporciona el conjunto de las funciones lineales. Sin embargo, aquí razonamos de manera intrínseca. Por los resultados generales sobre la composición de las aplicaciones lineales se sabe que F constituye un espacio vectorial cuando lo dotamos de las operaciones siguientes:

Suma.—Sean f^1 y f^2 dos funciones lineales; se les asocia la función f que hace corresponder al vector $\mathbf{v}\in E$ el escalar

$$f(\mathbf{v})=f^1(\mathbf{v})+f^2(\mathbf{v}).$$

Por definición, f es la suma de f^1 y f^2 y se escribe $f=f^1+f^2$. Esta operación confiere estructura de grupo al conjunto F, siendo el elemento neutro la aplicación nula.

Multiplicación por un escalar.—A la aplicación f y al escalar λ se les asocia $f'=\lambda f$ que, a todo $\mathbf{v}\in E$, hace corresponder el escalar $f'(\mathbf{v})=\lambda f(\mathbf{v})$.

El lector comprobará que se cumplen todos los axiomas de la estructura de espacio vectorial.

TEOREMA.—*La adición y la multiplicación por un escalar convierten en espacio vectorial el conjunto de las funciones lineales definidas sobre un espacio* E.

Se puede, pues, definir aplicaciones dependientes o independientes. Así, p. ej., f^1, f^2, f^3 son dependientes si existen tres escalares no todos nulos λ_1, λ_2, λ_3 tales que $\lambda_1 f^1 + \lambda_2 f^2 + \lambda_3 f^3$ sea la aplicación nula; es decir,

$$\lambda_1 f^1(\mathbf{v}) + \lambda_2 f^2(\mathbf{v}) + \lambda_3 f^3(\mathbf{v}) = 0$$

cualquiera que sea $\mathbf{v}$.

La representación analítica en una base permite efectuar las operaciones anteriores. Utilicemos la notación matricial; sean

$$f^1(\mathbf{v}) = \mathrm{AX} \qquad \text{y} \qquad f^2(\mathbf{v}) = \mathrm{BX}$$

dos funciones dadas. Se obtiene inmediatamente

$$f^1(\mathbf{v}) + f^2(\mathbf{v}) = (\mathrm{A} + \mathrm{B})\mathrm{X}$$

y la matriz (fila) correspondiente a $f^1 + f^2$ es $\mathrm{A} + \mathrm{B}$. Igualmente, la matriz fila de la forma λf^1 es $\lambda \mathrm{A}$.

16. Dimensión del espacio F. Base.—Estudiemos la dimensión del espacio F. Consideremos para ello una base cualquiera $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de E y asociemos a esta base las siguientes n funciones lineales: f^1, que toma el valor 1 para $\mathbf{e}_1$ y el valor 0 para los otros vectores de la base; en general, f^i (i de 1 a n) toma el valor 1 para $\mathbf{e}_i$ y el valor 0 para los restantes vectores de la base.

Las funciones f^1, f^2, ..., f^n constituyen un sistema de generadores de F, porque una función cualquiera f, que a los vectores $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ asocia los escalares a_1, a_2, ..., a_n puede representarse por la combinación lineal

$$f = a_1 f^1 + a_2 f^2 + \ldots + a_n f^n.$$

Las funciones f^1, f^2, ..., f^n son, además, independientes, porque si hubiera una relación de dependencia

$$\lambda_1 f^1 + \lambda_2 f^2 + \ldots + \lambda_n f^n = 0$$

debería verificarse para todo $\mathbf{v}$; eligiendo entonces $\mathbf{v} = \mathbf{e}_1$, de donde $f^1(\mathbf{e}_1) = 1$, $f^2(\mathbf{e}_1) = \ldots = f^n(\mathbf{e}_1) = 0$, se deduciría que $\lambda_1 = 0$. De manera análoga,

$$\lambda_2 = \lambda_3 = \ldots \lambda_n = 0,$$

lo que prueba la independencia.

La base $f^1, f^2, \ldots, f^n$ se llama base *dual* de $\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n$.

TEOREMA.—*El espacio* F *de las funciones lineales definidas en un espacio de n dimensiones es también de n dimensiones. A toda base de* E *se puede asociar una base de* F *obtenida considerando las n funciones* f^i *tales que* $f^i(\mathbf{e}_j) = \delta^i_j$ *(símbolo de Kronecker).*

Si E y F se hallan referidos a bases duales, al vector $\mathbf{v} = \sum x^i \mathbf{e}_i$ y a la función $f = \sum a_i f^i$ está asociado el escalar

$$f(\mathbf{v}) = a_1 x^1 + a_2 x^2 + \ldots + a_n x^n.$$

17. Dualidad.—A partir de E hemos construido el espacio F de las funciones lineales definidas sobre E. Sean $\mathbf{v} \in E$ y $f \in F$, ambos elegidos arbitrariamente. Al par $(\mathbf{v}, f)$ se asocia el escalar $f(\mathbf{v})$, valor de la función f para el vector $\mathbf{v}$. Considerando que $\mathbf{v}$ y f son vectores, diremos que $f(\mathbf{v})$ es su *producto escalar* y empleamos la notación simétrica $\langle \mathbf{v}, f \rangle$ para designar $f(\mathbf{v})$. El producto escalar es, por consiguiente, una aplicación del conjunto $E \times F$ en el cuerpo K. Esta aplicación es lineal con relación a $\mathbf{v}$, porque

$$\langle \lambda \mathbf{v} + \mu \mathbf{v}', f \rangle = \lambda \langle \mathbf{v}, f \rangle + \mu \langle \mathbf{v}', f \rangle \qquad [1]$$

en virtud de la definición de una función f, y lineal con relación a f, porque

$$\langle \mathbf{v}, \lambda f + \mu f' \rangle = \lambda \langle \mathbf{v}, f \rangle + \mu \langle \mathbf{v}, f' \rangle \qquad [2]$$

en virtud de las leyes de composición sobre F.

Estas propiedades establecen cierta reciprocidad entre E y F: todo vector de cada uno de estos espacios define una función lineal de los elementos del otro. Para F se puede decir más: todo $f \in F$ *es* una función lineal definida sobre E, y recíprocamente. Pero no sucede lo mismo para el espacio inicial E, donde los vectores son dados *a priori* y no están definidos como funciones lineales sobre F. El espacio Φ de estas funciones se construye aplicando a F el estudio hecho anteriormente para E. Los dos espacios E y Φ resultan, por consiguiente, independientes; sin embargo, hemos visto que a todo $\mathbf{v} \in E$ se asocia la función $\varphi \in \Phi$ tal que para todo $f \in F$ se tenga:

$$\langle \mathbf{v}, f \rangle = \varphi(f).$$

La aplicación de E en Φ así definida es lineal según [1], y es fácil ver

que es biyectiva [1]: los vectores de E representan de manera biunívoca las funciones lineales definidas sobre F. Pero habiendo definido los elementos de Φ por esta propiedad, nada se pierde, para estudiarlos, si el espacio Φ se asimila con el E: todo $\mathbf{v} \in E$ adquiere así la calidad de función lineal definida sobre F, lo que precisa la reciprocidad entre E y F y justifica el término de *espacio dual* empleado generalmente para designar F.

EJERCICIOS:

1. El espacio R^3 está referido a su base canónica. Determínese (por sus coordenadas en la base dual) la función lineal que a los vectores $\mathbf{v}_1=(2, 1, 0)$, $\mathbf{v}_2=(1, 1, 1)$, $\mathbf{v}_3=(1, 0, 2)$ asocia los escalares, respectivos, -1, 4, 2.

2. En R^3 se reemplaza la base canónica por la base $\mathbf{e}_1=(5, -1, 2)$, $\mathbf{e}_2=(4, 1, 1)$, $\mathbf{e}_3=(-3, 0, 1)$. Hállese la nueva expresión de la función lineal que, en la base canónica, se escribe $f(\mathbf{v})=2x^1+3x^2-x^3$.

3. Si n funciones lineales de un espacio E de n dimensiones son simultáneamente nulas para un mismo vector $\mathbf{v} \neq \mathbf{0}$, demuéstrese que estas funciones son dependientes.

4. Para que p funciones lineales definidas sobre un mismo espacio E de n dimensiones sean independientes es necesario y suficiente que exista al menos un vector de E para el cual las funciones tomen valores arbitrariamente elegidos.

5. Sea $\{\mathbf{e}_i\}$ una base de E. Estúdiese cómo se modifica la base dual $\{f^i\}$ de F cuando se someten los vectores de la base $\{\mathbf{e}_i\}$ a operaciones elementales (véase pág. 124).

18. Subespacios conjugados.—Sea E' un subespacio de E. Hallemos el conjunto de los vectores de F que tengan producto escalar nulo con un vector cualquiera de E'. Los papeles de los espacios E y F pueden, además, intercambiarse.

TEOREMA.—*El conjunto de los vectores de* F *(resp.* E*) cuyo producto escalar con todo vector de un subespacio* E' *de* E *(resp.* F' *de* F*)*

---

[1] Es inyectiva, porque los vectores $\mathbf{v}$ a los que corresponde la función $\varphi=0$ de Φ satisfacen a

$$\langle \mathbf{v}, f \rangle = 0, \qquad (\forall f).$$

Puesto que f es arbitraria, ello exige $\mathbf{v}=\mathbf{O}$ (para $\mathbf{v} \neq \mathbf{O}$ existen siempre funciones f tales que $f(\mathbf{v}) \neq 0$). El núcleo de la aplicación de E en Φ queda, por tanto, reducido al vector nulo. Como los espacios E y Φ tienen la misma dimensión (la de F), toda aplicación inyectiva es también biyectiva (véase página 344).

es nulo, constituye un subespacio F′ *de* F *(resp.* E′ *de* E). *La suma de las dimensiones de los dos subespacios es igual a n.*

Sea $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p\}$ una base de E′ completada mediante una base $\{\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n\}$ de un subespacio suplementario, y sea $\{f^1, f^2, \ldots, f^n\}$ la base dual de F. Si $f = \lambda_1 f^1 + \ldots + \lambda_n f^n$ es una cualquiera de las funciones buscadas, se debe tener

$$f(\mathbf{e}_1) = f(\mathbf{e}_2) = \ldots = f(\mathbf{e}_p) = 0.$$

Pero $f(\mathbf{e}_1) = \lambda_1 f^1(\mathbf{e}_1)$, porque $f^i(\mathbf{e}_1) = 0$ para $i \geqslant 2$. Es necesario, pues, elegir $\lambda_1 = 0$; igualmente, $\lambda_2 = \ldots = \lambda_p = 0$. Se obtiene así el conjunto de las funciones $f = \lambda_{p+1} f^{p+1} + \ldots + \lambda_n f^n$, que verifican las condiciones dadas y forman un subespacio F′ de $n - p$ dimensiones de F. La suma de las dimensiones es, evidentemente, n. El lector comprobará fácilmente que sirve la misma demostración si se permutan los papeles de E y F. De ello resulta la reciprocidad de E′ y F′: cada uno de estos subespacios constituye el conjunto de los vectores cuyo producto escalar con los vectores del otro es nulo. Estos dos espacios se llaman *conjugados.*

19. Determinación práctica.—Supongamos E y F referidos a dos bases determinadas, duales una de la otra $\{\mathbf{e}_i\}$ y $\{\mathbf{f}^i\}$. Uno de los subespacios, F′, p. ej., viene dado por un sistema de generadores

$$g^i = \sum_\alpha a^i_\alpha f^\alpha \qquad (i \text{ de } 1 \text{ a } p).$$

Se desea determinar un sistema de generadores del subespacio conjugado E′. Si $\mathbf{v} = \sum x^\alpha \mathbf{e}_\alpha$ es un vector cualquiera de E′, tenemos p relaciones

$$\left\{\begin{array}{l} a^1_1 x^1 + a^1_2 x^2 + \ldots + a^1_n x^n = 0, \\ a^2_1 x^1 + a^2_2 x^2 + \ldots + a^2_n x^n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a^p_1 x^1 + a^p_2 x^2 + \ldots + a^p_n x^n = 0, \end{array}\right. \qquad [\mathrm{I}]$$

que constituyen un sistema lineal y homogéneo con n incógnitas. El conjunto de las soluciones forma el subespacio E′. Si se reemplaza el sistema $\{g^i\}$ por un sistema equivalente, no cambian F′ ni E′. En estas condiciones, es posible reducir el sistema $\{g^i\}$ de manera que se obtengan r vectores independientes cuya matriz de coordenadas puede suponerse regularmente escalonada (véase pág. 142). El sistema lineal

obtenido [II] es equivalente al [I], y se escribe:

$$\left\{\begin{array}{l} a_1^1x^1+a_2^1x^2+\ldots+a_r^1x^r+a_{r+1}^1x^{r+1}+\ldots+a_n^1x^n=0,\\ \qquad b_2^2x^2+\ldots+b_r^2x^r+b_{r+1}^2x^{r+1}+\ldots+b_n^2x^n=0,\\ \qquad\qquad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \qquad\qquad\qquad h_r^rx^r+h_{r+1}^rx^{r+1}+\ldots+h_n^rx^n=0. \end{array}\right. \qquad [II]$$

Los coeficientes a_1^1 b_2^2, ..., h_r^r son todos diferentes de 0. La resolución es inmediata: se dan valores arbitrarios a las incógnitas x^{r+1}, ..., x^n y las ecuaciones sucesivas (yendo hacia arriba) proporcionan las restantes, x^r, x^{r-1}, ..., x^1. La elección arbitraria de las $n-r$ últimas incógnitas permite formar $n-r$ soluciones independientes, que son:

$$\begin{array}{l} \mathbf{v}_1=(x_1^1, x_1^2, \ldots, x_1^r, 1, 0, \ldots, 0),\\ \mathbf{v}_2=(x_2^1, x_2^2, \ldots, x_2^r, 0, 1, \ldots, 0),\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \mathbf{v}_{n-r}=(x_{n-r}^1, x_{n-r}^2, \ldots, x_{n-r}^r, 0, 0, \ldots, 1). \end{array}$$

Las $n-r$ últimas columnas muestran claramente la independencia. Tenemos, por tanto, una base de E'.

OBSERVACIONES.—1. Cabe decir que el sistema [I] (reducido a r ecuaciones independientes) constituye las ecuaciones del subespacio E'. Quedan así generalizadas las nociones clásicas de geometría analítica de 2 ó 3 dimensiones.

2. Recíprocamente, si se da E' mediante un sistema de generadores, será posible hallar sus ecuaciones, es decir, las condiciones necesarias y suficientes para que el vector $\sum x^i\mathbf{e}_i$ pertenezca al mismo. Los papeles recíprocos de E' y F' prueban que este problema se resuelve de la misma manera. Basta cambiar los papeles de las a y de las x.

EJERCICIOS:

1. Determínese una base del conjugado del subespacio E' de R^4 engendrado por (1, 1, 0, 1), (0, 1, 2, 0), (0, 0, 3, 4). Dedúzcase de ello la ecuación de este subespacio.

2. Determínese, mediante un sistema de generadores, el subespacio de R^4 que tiene por ecuaciones

$$2x_1+x_2-x_3+2x_4=0, \quad x_1-2x_2+x_3+x_4=0, \quad x_1+x_2+2x_3-x_4=0.$$

3. A todo subespacio E', E'', ... de E, asociamos el subespacio conjugado F', F'', ... de F. Pruébese que la correspondencia verifica las propiedades siguientes:

a) si $E' \subseteq E''$, tenemos $F'' \subseteq F'$;
b) si $E_1 = E' + E''$, tenemos $F_1 = F' \cap F''$;
c) si $E_1 = E' \cap E''$, tenemos $F_1 = F' + F''$.

4. Si $E = E' \oplus E''$ es una descomposición de E en suma directa, demuéstrese que los subespacios conjugados F' y F'' dan una descomposición de F en suma directa.

ESPACIOS DUALES

20. Dualidad.—El estudio precedente no permite establecer una reciprocidad completa entre los espacios E y F: la existencia de E se fija *a priori*, y de ello se deduce la existencia y definición de F. Consideremos de nuevo la teoría de la dualidad, pero de modo que E y E* desempeñen idénticos papeles. Los definimos independientemente uno del otro y no hacemos ninguna hipótesis particular sobre los vectores que los constituyen.

Definición.— Dos espacios vectoriales E y E*, sobre un mismo cuerpo K, se llaman *duales* uno del otro si verifican las condiciones siguientes:

CONDICIÓN I.—A todo $v \in E$ y a todo[1] $v^* \in E^*$ sabemos asociar un escalar K, llamado *producto escalar*, y denotado por $\langle v, v^* \rangle$.
Este producto es lineal con relación a v:

$$\langle \lambda v + \mu v', v^* \rangle = \lambda \langle v, v^* \rangle + \mu \langle v', v^* \rangle$$

y lineal con relación a v^*:

$$\langle v, \lambda v^* + \mu v'^* \rangle = \lambda \langle v, v^* \rangle + \mu \langle v, v'^* \rangle.$$

Dos vectores cuyo producto escalar es nulo se dicen *ortogonales*. De la doble linealidad resulta que el vector nulo de E (o de E*) es ortogonal a todo vector de E* (o de E).

CONDICIÓN II.—Se supone además que el vector nulo de E es el único que es ortogonal a todo vector de E*, y que el vector nulo de E* es el único que es ortogonal a cualquier vector de E.

CONSECUENCIAS.—Sean F (resp. F*) los espacios de las funciones lineales definidas sobre E (resp. E*). Consideremos un elemento de-

[1] Designamos por v, w, ... los elementos de E, y por v^*, w^*, ... los de E*, sin suponer ninguna relación entre v y v^*, p. ej. Igualmente f y f^* son funciones lineales definidas, respectivamente, sobre E y E*.

terminado $\mathbf{v}$ de E. El producto escalar $\langle \mathbf{v}, \mathbf{w}^* \rangle$, en el que $\mathbf{w}^*$ describe E^*, es una función lineal de este vector. Existe, pues, en el espacio F^* un elemento bien determinado f^* tal que para $\mathbf{w}^*$ arbitrario, se verifica que

$$\langle \mathbf{v}, \mathbf{w}^* \rangle = f^*(\mathbf{w}^*).$$

Igualmente a $\mathbf{v}^* \in \mathrm{E}^*$ corresponde $f \in \mathrm{F}$ tal que, para $\mathbf{w}$ arbitrario, se tiene:

$$\langle \mathbf{w}, \mathbf{v}^* \rangle = f(\mathbf{w}).$$

Queda, pues, definida de manera intrínseca una aplicación de cada uno de los espacios E (o E^*) en el espacio de las funciones lineales definidas sobre el otro.

Se comprueba inmediatamente que estas aplicaciones son lineales; si a $\mathbf{v}$ y a $\mathbf{w}$ corresponden, respectivamente, f^* y g^*, a $\lambda\mathbf{v} + \mu\mathbf{w}$ corresponde una función h^* tal que, para $\mathbf{v}^*$ arbitrario, se tiene:

$$\langle \lambda\mathbf{v} + \mu\mathbf{w}, \mathbf{v}^* \rangle = h^*(\mathbf{v}^*).$$

Pero

$$\langle \lambda\mathbf{v} + \mu\mathbf{w}, \mathbf{v}^* \rangle = \lambda\langle \mathbf{v}, \mathbf{v}^* \rangle + \mu\langle \mathbf{w}, \mathbf{v}^* \rangle = \lambda f^*(\mathbf{v}^*) + \mu g^*(\mathbf{v}^*)$$

de donde

$$h^* = \lambda f^* + \mu g^*.$$

La condición II nos va a permitir precisar las propiedades de estas aplicaciones.

TEOREMA 1.—*Dos espacios duales* E *y* E^* *de dimensión finita tienen la misma dimensión.*

Sea m la dimensión de E (por tanto, de F); n la de E^* (por tanto, de F^*). Demostremos que la aplicación de E en F^* definida anteriormente es inyectiva. Si $\mathbf{v}$ es un elemento del núcleo, le corresponde la función nula de F^* y, por consiguiente,

$$\langle \mathbf{v}, \mathbf{v}^* \rangle = 0,$$

para todo $\mathbf{v}^*$. Esto equivale a decir que $\mathbf{v}$ es ortogonal a todos los vectores de E^* y la condición II exige que se elija $\mathbf{v} = \mathbf{0}$. Como la aplicación es inyectiva, se conserva la dimensión. Su rango es, pues, igual a m, dimensión de E, lo que exige $m \leqslant n$.

Pero puede repetirse el razonamiento cambiando los papeles de E y E^*. Se deduce entonces que $n \leqslant m$; por consiguiente, $m = n$.

TEOREMA 2.—*Cada uno de los espacios* E *o* E* *es naturalmente isomorfo al espacio de las funciones lineales definidas sobre el otro.*

Del teorema 1 resulta que E y F* tienen la misma dimensión, y la aplicación lineal de E en F*, ya considerada, no solo es inyectiva, sino también biyectiva (véase pág. 344). Se trata, por tanto, de un isomorfismo natural, ya que queda establecido con independencia de la elección de las bases.

Dicho de otra manera, al conjunto de los $\mathbf{v} \in E$ corresponde de manera biunívoca (y lineal) el conjunto de las funciones definidas sobre E*; una función f^* dada se representa de manera única por el producto $\langle \mathbf{v}, \mathbf{v}^* \rangle$, en donde $\mathbf{v}$ está asociado a f^*.

EJEMPLOS.—1. Un espacio E de dimensión finita y el espacio F de las funciones lineales definidas sobre E forman un par de espacios duales. El producto escalar es $\langle \mathbf{v}, f \rangle = f(\mathbf{v})$ y se verifica también la condición II. En efecto: $f(\mathbf{v}) = 0$, supuesta cierta para $\mathbf{v}$ cualquiera, implica $f = O$; si se supone cierta para un $\mathbf{v}$ determinado y f arbitraria, exige que $\mathbf{v} = \mathbf{0}$, porque para $\mathbf{v} \neq \mathbf{0}$ se puede hallar siempre (teorema de la pág. 347) una función no nula para este vector.

De ello resulta que E es naturalmente isomorfo al espacio de las funciones lineales definidas sobre F: esta es la propiedad vista en la página 350.

2. Veremos en el capítulo 19 que un espacio dado de dimensión finita puede ser su propio dual: $E = E^*$.

OBSERVACIONES.—Ciertos resultados de la sección precedente pueden traducirse utilizando esta terminología.

1. Existe un vector y uno solo de E* (o de E) que tenga un producto escalar dado con cada uno de los n vectores de una base de E (o de E*). En el isomorfismo natural, este vector de E* corresponde a la función f obtenida aplicando el teorema de la página 347.

2. A todo subespacio de uno de los espacios corresponde un subespacio ortogonal [1] del otro. Existe reciprocidad entre estos dos subespacios y la suma de sus dimensiones es igual a n.

21. Bases duales.—Sean E y E* dos espacios duales, y $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ una base de E. De acuerdo con la primera observación anterior, existe un vector y uno solo $\mathbf{e}^{*1}$ de E*, tal que

$$\langle \mathbf{e}_1, \mathbf{e}^{*1} \rangle = 1, \qquad \langle \mathbf{e}_i, \mathbf{e}^{*1} \rangle = 0 \qquad \text{para} \qquad i \geqslant 2.$$

[1] Decimos ortogonal en lugar de conjugado. Todo vector de cada subespacio es ortogonal a todo vector del otro.

Se puede generalizar esto y definir n vectores $\mathbf{e}^{*i}$ por las condiciones

$$\langle \mathbf{e}_i, \mathbf{e}^{*j} \rangle = \delta_i^j \qquad (0 \text{ si } i \neq j,\ 1 \text{ si } i = j). \tag{1}$$

Se obtiene así una base [1] $\{\mathbf{e}^{*1}, \mathbf{e}^{*2}, \ldots, \mathbf{e}^{*n}\}$ de $\mathbf{E}^*$, que se llama base dual de $\{\mathbf{e}_i\}$. La relación general [1] demuestra la reciprocidad entre estas dos bases.

TEOREMA 1.—*La coordenada de orden i de un vector de* E *(resp.* $\mathbf{E}^*$*) en una base es igual al producto escalar de este vector por el vector de orden i de la base dual de* $\mathbf{E}^*$ *(resp.* E*).*

Partiendo de $\mathbf{v} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$ y multiplicando escalarmente por $\mathbf{e}^{*i}$, se obtiene, en virtud de [1],

$$\langle \mathbf{v}, \mathbf{e}^{*i} \rangle = x^i.$$

TEOREMA 2.—*Si dos espacios duales están referidos a bases duales, el producto escalar de los vectores*

$$\mathbf{v} = \sum x^\alpha \mathbf{e}_\alpha \qquad y \qquad \mathbf{v}^* = \sum y_\beta \mathbf{e}^{*\beta}$$

tiene por expresión [2]

$$\langle \mathbf{v}, \mathbf{v}^* \rangle = \sum x^\alpha y_\alpha.$$

La bilinealidad permite escribir:

$$\langle \mathbf{v}, \mathbf{v}^* \rangle = \sum_{\alpha,\beta} x^\alpha y_\beta \langle \mathbf{e}_\alpha, \mathbf{e}^{*\beta} \rangle = \sum_{\alpha,\beta} x^\alpha y_\beta \delta_\alpha^\beta$$

y el significado de δ_α^β demuestra que en la doble sumación se introducen solo los términos correspondientes a dos índices α y β iguales.

Es cómodo adoptar una escritura matricial. Si a $\mathbf{v}$ corresponde la matriz columna $\mathrm{X} = [x^i]$, se elegirá para $\mathbf{v}^*$ una matriz fila $\mathrm{Y} = [y_i]$, de manera que

$$\langle \mathbf{v}, \mathbf{v}^* \rangle = \mathrm{YX}.$$

---

[1] Una relación de dependencia $\lambda_1\mathbf{e}^{*1} + \ldots + \lambda_n\mathbf{e}^{*n} = 0$ implicaría $\lambda_1 = 0$ si se multiplica escalarmente por $\mathbf{e}_1$; $\lambda_2 = 0$ si se multiplica por $\mathbf{e}_2$, etc.

[2] Nos ajustamos al convenio corriente: sumación sobre un índice repetido arriba y abajo.

22. Cambios de base.—Consideremos en E el cambio de base definido por las fórmulas

$$\mathbf{f}_i = \sum a_i^\alpha \mathbf{e}_\alpha, \qquad [1]$$

o, recíprocamente,

$$\mathbf{e}_i = \sum b_i^\alpha \mathbf{f}_\alpha \qquad [1\ bis]$$

A cada base de E corresponde de manera única una base dual de E*:

$$\langle \mathbf{e}_i, \mathbf{e}^{*j} \rangle = \delta_i^j, \qquad [2]$$

$$\langle \mathbf{f}_i, \mathbf{f}^{*j} \rangle = \delta_i^j \qquad [2\ bis]$$

y el cambio en E define un cambio de base en E*, que debemos caracterizar. Se hace

$$\mathbf{f}^{*i} = \sum a_\alpha'^i \mathbf{e}^{*\alpha}, \qquad [3]$$

y se hallan los coeficientes $a_j'^i$. Por el teorema 1 del apartado anterior, se tiene:

$$a_j'^i = \langle \mathbf{f}^{*i}, \mathbf{e}_j \rangle.$$

Reemplazando los $\mathbf{e}_j$ con ayuda de las $\mathbf{f}$ (relaciones [1 *bis*]) y utilizando las condiciones de ortogonalidad [2 *bis*], se obtiene inmediatamente,

$$a_j'^i = b_j^i,$$

Por tanto, [3] se escribe:

$$\mathbf{f}^{*i} = \sum b_\alpha^i \mathbf{e}^{*\alpha}, \qquad [3]$$

o

$$\mathbf{e}^{*i} = \sum a_\alpha^i \mathbf{f}^{*\alpha}. \qquad [3\ bis]$$

Es cómodo interpretar este resultado con ayuda de las matrices. Si $P = [a_i^j]$ [1] es la matriz del cambio [1], la matriz del cambio [3] será $[b_j^i] = P^{-1}$. Solo que la posición del índice de sumación en [3] nos enseña que esta nueva matriz de cambio debe leerse de acuerdo con el

[1] En la igualdad $P = [a_i^j]$ se debe tomar i como índice de columna y j como índice de fila si queremos que los componentes de los nuevos vectores de la base se lean en las columnas.

convenio adoptado anteriormente para la escritura matricial de los vectores de E^*: los componentes de los nuevos vectores de la base se leen en las filas.

TEOREMA.—*Si* P *es la matriz de cambio de una base a otra de* E, *la matriz de cambio para las bases duales de* E^* *es* P^{-1}.

En estas condiciones, es fácil establecer las fórmulas de transformación de coordenadas para un vector.

Para E: Sea v de matriz columna X en la base antigua y X′ en la nueva; sabemos que

$$X = PX', \qquad \text{o bien,} \qquad X' = P^{-1}X.$$

Para E^*: Sea v^* de matriz fila Y en la antigua base e Y′ en la nueva. Tenemos:

$$v^* = \sum y_\alpha e^{*\alpha} = \sum_{\alpha,\beta} y_\alpha a_\beta^\alpha f^{*\beta} = \sum_\beta y'_\beta f^{*\beta},$$

de donde,

$$y'_i = \sum a_i^\alpha y_\alpha.$$

Escribiendo el segundo miembro en la forma

$$\sum y_\alpha a_i^\alpha,$$

se identifica inmediatamente un producto de matrices:

$$Y' = YP.$$

OBSERVACIÓN.—Las fórmulas

$$X = PX' \qquad \text{e} \qquad Y' = YP$$

prueban que las coordenadas de los vectores de E o de E^* no se comportan de igual modo respecto de los cambios de base de E.

Para E, las coordenadas $(x^1, x^2, \ldots, x^n)$ de un vector se llaman *contravariantes* con relación a las transformaciones de los vectores de la base. Por abuso de lenguaje se dice que el propio vector v es *contravariante* (cuando es sencillamente invariante). Al contrario, las componentes $(y_1, y_2, \ldots, y_n)$ de un vector de E^* se denominan *covariantes* con relación a las transformaciones de los vectores de la base de E, y el vector mismo se llama *covariante*. Naturalmente, la contravarian-

cia y la covariancia están definidas solo con relación a uno de los espacios, al cual se atribuye cierta prioridad respecto al otro.

La distinción entre estas dos especies de vectores es la razón esencial para elegir la posición de los índices (x^i o y_j) en la escritura de las coordenadas.

APLICACIONES TRASPUESTAS

23. Traspuesta de una aplicación lineal.—Sean E y F dos espacios vectoriales sobre el mismo cuerpo K, E* un dual de E, y F* un dual de F. Llamemos $\mathbf{v}$, $\mathbf{v}_1$, $\mathbf{v}_2$, ... y $\mathbf{v}^*$, $\mathbf{v}_1^*$, ... a los elementos de E y de E*; $\mathbf{w}$, $\mathbf{w}_1$, ... y $\mathbf{w}^*$, $\mathbf{w}_1^*$, ... a los vectores de F y de F*, respectivamente.

Vamos a asociar a toda aplicación lineal T de E en F una aplicación de F* en E* llamada *traspuesta* de T y denotada por ${}^tT$. Para definirla, debemos asociar a todo vector $\mathbf{w}^* \in F^*$ un vector $\mathbf{v}^* \in E^*$, siendo la aplicación $\mathbf{w}^* \longrightarrow \mathbf{v}^*$ lineal. Sabemos que el vector $\mathbf{w}^*$ define una función lineal φ de los elementos de F, de modo que, para $\mathbf{w}$ cualquiera, se tenga:

$$\varphi(\mathbf{w}) = \langle \mathbf{w}, \mathbf{w}^* \rangle.$$

Observemos que es posible componer las aplicaciones T y φ, puesto que T se efectúa de E en F, y φ, de F en K. La aplicación lineal obtenida $f = \varphi \circ T$, se efectúa de E en K; por consiguiente, es una función lineal definida sobre E y tal que, para todo $\mathbf{v}$,

$$f(\mathbf{v}) = \varphi(T\mathbf{v}). \qquad [1]$$

Es sabido que a dicha función corresponde, de manera intrínseca, un vector $\mathbf{v}^* \in E^*$ tal que

$$f(\mathbf{v}) = \langle \mathbf{v}, \mathbf{v}^* \rangle.$$

Escribimos:

$$\mathbf{v}^* = {}^tT\mathbf{w}^*$$

y vamos a estudiar la aplicación traspuesta ${}^tT$ así definida.

Propiedad fundamental.—*Si* $\mathbf{v}$ *es un vector de* E *y* $\mathbf{w}^*$ *un vector de* F*, los transformados* $T\mathbf{v}$ *y* ${}^tT\mathbf{w}^*$ *en la aplicación* T *y en su traspuesta* ${}^tT$

verifican la condición:

$$\langle v, {}^tTw^* \rangle = \langle Tv, w^* \rangle. \qquad [1]$$

Se trata de una traducción inmediata de la definición de la aplicación ${}^tT$. En efecto, con las mismas notaciones que anteriormente, el segundo miembro se escribe $\varphi(Tv)$ y el primero $f(v)$. La igualdad a demostrar equivale, pues, a [1].

TEOREMA.—*La traspuesta es una aplicación lineal. La operación es recíproca:* ${}^t({}^tT) = T$.

En [1] hacemos $w^* = w_1^* + w_2^*$, de donde

$$\langle v, {}^tT(w_1^* + w_2^*) \rangle = \langle Tv, w_1^* + w_2^* \rangle = \langle Tv, w_1^* \rangle + \langle Tv, w_2^* \rangle$$
$$= \langle v, {}^tTw_1^* \rangle + \langle v, {}^tTw_2^* \rangle.$$

Como la relación obtenida se verifica cualquiera que sea v, evidentemente se tiene:

$${}^tT(w_1^* + w_2^*) = {}^tTw_1^* + {}^tTw_2^*.$$

La reciprocidad se obtiene también a partir de [1]. Si T' es la transformación de E en F traspuesta de ${}^tT$, se tendrá que

$$\langle w^*, T'v \rangle = \langle v, {}^tTw^* \rangle.$$

Comparando con [1] resulta que $\langle Tv, w^* \rangle = \langle T'v, w^* \rangle$, de donde $Tv = T'v$ y $T = T'$ (w^* y v son arbitrarios).

24. Traspuestas de una suma y de un producto. TEOREMA 1.—*La traspuesta de la suma de dos aplicaciones lineales* T *y* T_1 *es la suma de las traspuestas.*

Se escriben las igualdades [1] para T y T_1 eligiendo los mismos vectores v y w^*. Sumando y utilizando la linealidad de las aplicaciones y del producto escalar se deduce inmediatamente la igualdad buscada.

De la misma manera se obtiene la fórmula

$${}^t(\lambda T + \mu T_1) = \lambda\,{}^tT + \mu\,{}^tT_1.$$

TEOREMA 2.—*La traspuesta de la aplicación compuesta* $T_2 \circ T_1$ *es la compuesta de las aplicaciones traspuestas consideradas en sentido inverso:*

$${}^t(T_2 \circ T_1) = {}^tT_1 \circ {}^tT_2.$$

Supongamos que T_1 aplique E en F, y T_2, F en G; sean

$$w = T_1 v \qquad y \qquad z = T_2 w = (T_2 \circ T_1)v.$$

Igualmente, para los espacios duales,

$$w^* = {}^tT_2 z^* \qquad y \qquad v^* = {}^tT_1 w^* = ({}^tT_1 \circ {}^tT_2)z^*.$$

Pero la igualdad [1] se escribe con este sistema de notaciones:

para T_1:

$$\langle v, v^* \rangle = \langle w, w^* \rangle;$$

para T_2:

$$\langle w, w^* \rangle = \langle z, z^* \rangle.$$

Comparando y reemplazando v^* y z, se obtiene la igualdad

$$\langle v, {}^tT_1 \circ {}^tT_2 z^* \rangle = \langle T_2 \circ T_1 v, z^* \rangle,$$

que expresa el resultado deseado.

25. Núcleo y rango de la aplicación traspuesta.—El núcleo es el conjunto de los vectores w^* tales que $v^* = {}^tTw^*$ sea el vector nulo de E^*. Para tal vector, la igualdad [1] se convierte en

$$\langle v, {}^tTw^* \rangle = \langle Tv, w^* \rangle = 0;$$

la condición $\langle Tv, w^* \rangle = 0$, supuesta verificada para un v arbitrario, caracteriza a los elementos w^* del núcleo. Cuando v describe E, Tv describe el subespacio imagen T(E) de F; la condición anterior expresa que w^* es ortogonal a este subespacio. Describe el subespacio ortogonal. De donde el teorema siguiente.

TEOREMA.—*El núcleo de la aplicación traspuesta de* T *es el subespacio de* F^* *ortogonal al subespacio imagen,* T(E), *de la transformación dada.*

De ello se deduce una propiedad del rango.

TEOREMA.—*Una aplicación lineal y su traspuesta tienen el mismo rango.*

Si es n la dimensión común de E y de E^*, m la de F y F^* y r la dimensión de T(E), el núcleo de 'T tiene por dimensión $m - r$ (véase página 356, observación 2). La relación fundamental (pág. 343) da el

rango r_1 de ${}^tT$:

$$r_1 = m - (\text{dim. núcleo}) = m - (m - r) = r.$$

OBSERVACIÓN.—El teorema anterior conduce a un método teórico para expresar que un vector $\mathbf{w} \in F$ pertenece a T(E). Para ello es necesario y suficiente que sea ortogonal al núcleo de la transformación traspuesta. Si recordamos los resultados obtenidos en el estudio de las ecuaciones lineales (véase pág. 344), se sabe que $T\mathbf{x} = \mathbf{y}_0$ es posible si, y solo si, $\mathbf{y}_0 \in T(E)$. Si ${}^tT$ es la traspuesta de T, la ecuación

$${}^tT\mathbf{z} = \mathbf{O}$$

se llama ecuación homogénea traspuesta; admite por solución el núcleo de ${}^tT$. De ello, resulta:

TEOREMA.—*Para que sea posible una ecuación lineal* $T(\mathbf{x}) = \mathbf{y}_0$ *es necesario y suficiente que el vector* $\mathbf{y}_0$ *sea ortogonal a todas las soluciones de la ecuación homogénea traspuesta* ${}^tT\mathbf{z} = \mathbf{O}$.

EJERCICIOS:

1. Demuéstrese que, si T es un isomorfismo de E en F, la transformación ${}^tT$ es un isomorfismo de F^* en E^* y que $({}^tT)^{-1} = {}^t(T^{-1})$.

2. Si T es una aplicación lineal de E en un espacio dual E^*, ¿qué cabe decir de ${}^tT$?

3. Sean E y E^* dos espacios duales. Si es T una aplicación idempotente ($T^2 = T$) de E en sí mismo, y ${}^tT$ su traspuesta, demuéstrese que los subespacios imágenes T(E) y ${}^tT(E^*)$ son también duales uno del otro para el producto escalar definido entre E y E^*.

4. Sean dos espacios duales E y E^*; E' y E'' dos subespacios suplementarios de E, y E'* y E''* los subespacios respectivamente ortogonales de E^*. Pruébese que E' y E''* son duales entre sí, así como E'' y E'*.

5. Se considera el espacio E de las sucesiones de números reales tales que solo sean distintos de cero un número finito de términos (véase página 212). Caracterícese el espacio de las funciones lineales definidas sobre E.

CAPITULO XVI

APLICACIONES MULTILINEALES. DETERMINANTES

La definición de aplicaciones entre espacios vectoriales se extiende al caso de aplicaciones multilineales. El lector observará la profunda analogía existente entre todos estos desarrollos del álgebra lineal.

Las definiciones se expresan de manera intrínseca; es decir, independientemente de toda base de referencia de los vectores. Más adelante, se podrá utilizar una referencia particular para determinar prácticamente una aplicación.

No insistiremos en el estudio general de las aplicaciones multilineales, lo que nos llevaría al cálculo tensorial, pero vamos a precisar el caso de las aplicaciones alternadas. La expresión analítica en una base conduce entonces, de manera lógica, a un mecanismo de cálculo bastante complejo: el del determinante asociado a una matriz cuadrada.

Así justificada la definición, se prosigue el estudio de los determinantes en el resto del capítulo, pero independientemente de las dos primeras secciones, cuyo carácter, más bien abstracto, podría crear dificultades a algunos lectores.

APLICACIONES MULTILINEALES

1. Definición.—Sea E un espacio vectorial sobre un cuerpo K, y consideremos una aplicación del conjunto producto $(E)^p$ en K; es decir, una correspondencia que asocia a todo conjunto ordenado $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ formado por p vectores de E, un elemento bien determinado de K, $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$. La aplicación se llama multilineal cuando es lineal con relación a cada uno de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, o sea cuando conserva las combinaciones lineales efectuadas sobre cualquiera de ellos. Así, p. ej., si el i-ésimo vector $\mathbf{v}_i$ se reemplaza por $\lambda\mathbf{v}_i + \mu\mathbf{v}'_i$ se deberá verificar:

$$f(\mathbf{v}_1, \ldots, \lambda\mathbf{v}_i + \mu\mathbf{v}'_i, \ldots \mathbf{v}_p) = \lambda f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_p) + \\ + \mu f(\mathbf{v}_1, \ldots, \mathbf{v}'_i, \ldots, \mathbf{v}_p). \qquad [1]$$

En particular, la multiplicación de $\mathbf{v}_i$ por λ multiplica el valor de f por el mismo escalar. Como consecuencia inmediata, resulta que si uno de los vectores es nulo, la aplicación toma el valor nulo.

OBSERVACIONES. *a)* Para $p=1$ obtenemos de nuevo la definición de función lineal definida sobre E.

b) Es posible definir aplicaciones multilineales en circunstancias más generales; por una parte, puede ampliarse el dominio de definición, suponiendo que elegimos $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ en p espacios distintos, $E_1, E_2, \ldots, E_p$; por otra, en lugar de hacer la aplicación en K cabe hacerla en otro espacio vectorial F. Si el cuerpo de los coeficientes es el mismo, el lector comprobará fácilmente que la multilinealidad conserva su significado. Nos limitaremos aquí al caso examinado inicialmente, observando, sin embargo, que las definiciones de producto escalar (página 354) y de producto exterior (pág. 391) utilizan la definición generalizada.

c) Para recordar que la aplicación se hace en el cuerpo K diremos con frecuencia *función multilineal,* en lugar de aplicación.

Consecuencias.—Se puede utilizar el carácter lineal con relación a cada vector para transformar una función f. Así, p. ej.,

a) $$f(\lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2, \ldots, \lambda_p\mathbf{v}_p) = \lambda_1\lambda_2 \ldots \lambda_p f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p);$$

b) Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ son sumas de $k_1, k_2, \ldots, k_p$ vectores, respectivamente, se puede descomponer $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ en una suma de $k_1 \times k_2 \times \ldots \times k_p$ valores de la misma función.

c) En general, suponemos que los p vectores son combinaciones lineales de cierto número q de otros vectores:

$$\mathbf{v}_1 = \sum_1^q \lambda_{1i}^{\alpha_1}\mathbf{w}_{\alpha_1}, \quad \ldots, \quad \mathbf{v}_p = \sum_1^q \lambda_{1i}^{\alpha_p}\mathbf{w}_{\alpha_p}.$$

(Los índices mudos $\alpha_1, \alpha_2, \ldots, \alpha_p$ varían de 1 a q independientemente uno del otro.) Se puede transformar $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ con ayuda de las dos propiedades anteriores, lo que da

$$f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p) = \sum_{\alpha_1\alpha_2\ldots\alpha_p} \lambda_1^{\alpha_1}\lambda_2^{\alpha_2} \ldots \lambda_p^{\alpha_p} f(\mathbf{w}_{\alpha_1}, \mathbf{w}_{\alpha_2}, \ldots, \mathbf{w}_{\alpha_p}).$$

En el segundo miembro, aparece la suma de los q^p valores de f obtenidos al hacer variar los índices separadamente de 1 a q; estos valo-

res $f(w_{\alpha_1}, w_{\alpha_2}, \ldots, w_{\alpha_p})$ corresponden a las variaciones con repetición de los q vectores $w_1, w_2, \ldots, w_p$ tomados de p en p.

Operaciones.—Se define la suma de dos funciones, f^1 y f^2, p veces lineales, así como el producto de f^1 por $\lambda \in K$, de la misma forma que en el caso $p=1$ (véase pág. 348). El conjunto de estas funciones adquiere así la estructura de espacio vectorial.

2. Expresión analítica en una base.—La definición anterior posee carácter intrínseco. Limitándonos al caso de un espacio E de n dimensiones, se puede traducir $f(v_1, v_2, \ldots, v_p)$ analíticamente utilizando las expresiones de los vectores en una base $\{e_1, e_2, \ldots, e_n\}$. Poniendo

$$v_i = \sum x_i^\alpha e_\alpha \qquad (i \text{ de } 1 \text{ a } p)$$

y aplicando la propiedad *c)* anterior, se obtiene:

$$f(v_1, v_2, \ldots, v_p) = \sum_{\alpha_1, \alpha_2, \ldots, \alpha_p} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_p^{\alpha_p} f(e_{\alpha_1}, e_{\alpha_2}, \ldots, e_{\alpha_p}). \qquad [1]$$

En el segundo miembro, la suma se hace sobre todos los índices $\alpha_1, \alpha_2, \ldots, \alpha_p$, que varían independientemente de 1 a n (n^p términos). Haciendo

$$A_{\alpha_1 \alpha_2 \ldots \alpha_p} = f(e_{\alpha_1}, e_{\alpha_2}, \ldots, e_{\alpha_p}), \qquad [2]$$

los A son elementos bien determinados de K, independientes de la elección de los vectores, y se tiene:

$$f(v_1, v_2, \ldots, v_p) = \sum_{\alpha_1, \ldots, \alpha_n} A_{\alpha_1 \alpha_2 \ldots \alpha_n} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}. \qquad [1\ bis]$$

El segundo miembro es un polinomio homogéneo de grado p respecto al conjunto de las variables x_i^α (i de 1 a p). Cada término contiene una, y solo una, de las coordenadas de cada vector, y dicho segundo miembro es una forma lineal con relación a las coordenadas de un vector determinado. Se dice que es una forma p veces lineal.

TEOREMA.—*Toda función p veces lineal se expresa en una base cualquiera por una forma p veces lineal de las coordenadas de los vectores. Los coeficientes son los valores de la función para las variaciones con repetición de los vectores de la base tomados p a p.*

Recíprocamente, la elección, en una base determinada, de una forma p veces lineal de las coordenadas de p vectores arbitrarios determina una función p veces lineal, si convenimos en ligar intrínsecamente el valor así asociado a todo conjunto de p vectores.

Por supuesto, la expresión formal de la función así definida depende de la elección de la base. No estudiaremos aquí el mecanismo que permite relacionar entre sí los sistemas de coeficientes $A_{\alpha_1}, \ldots, \alpha_n$ cuando se hace un cambio de base [1].

Las funciones p veces lineales de E se llaman también *tensores p veces covariantes* [2] y el estudio de sus propiedades pertenece al *cálculo tensorial*. En el capítulo 17 damos algunas indicaciones sobre los tensores dos veces covariantes.

EJERCICIOS:

1. El espacio R^2 está referido a la base canónica e_1, e_2. Al par de vectores $v_1=(x_1^1, x_1^2)$, $v_2=(x_2^1, x_2^2)$ se asocia el escalar $f(v_1, v_2)=x_1^1x_2^2$. Determínese la nueva expresión de esta función cuando se toma como nueva base

$$e_1'=e_1+e_2,\quad e_2'=e_1-e_2.$$

2. La misma cuestión considerando la aplicación $f_1(v_1, v_2)=x_1^1x_2^2-x_1^2x_2^1$.

3. A todo par (v, w) de vectores de R^2 se asocia el vector z de R^4, siendo la aplicación bilineal. Se escribe $z=T(v, w)$. Si (e_1, e_2) es una base de R^2, demuéstrese que la aplicación queda completamente definida por

$$T(e_1, e_1)=f_1 \qquad T(e_1, e_2)=f_2, \qquad T(e_2, e_1)=f_3, \qquad T(e_2, e_2)=f_4,$$

donde f_1, f_2, f_3, f_4 pertenecen a R^4. Suponiendo, además, independientes estos cuatro vectores, caracterícese en R^4 el subconjunto de los vectores imágenes de los pares (v, w). Si $z=z^if_i$, se verá que un vector es imagen si, y solo si, $z^1z^4-z^2z^3=0$. ¿Es este subconjunto un subespacio de R^4?

APLICACIONES ANTISIMETRICAS

3. Transformada de una aplicación por una sustitución.—Recordemos que dados p elementos ordenados $(v_1, v_2, \ldots, v_p)$, el conjunto $\mathcal{S}$ de las $p!$ sustituciones que hacen pasar de la permutación considerada a una cualquiera de las otras constituye un grupo (véase pág. 60). Si $s \in \mathcal{S}$, representamos por $(sv_1, sv_2, \ldots, sv_p)$ la permutación imagen de $(v_1, v_2, \ldots, v_p)$ por s.

[1] Sin embargo, la fórmula [2] permite tratar la cuestión en casos sencillos.

[2] Con mayor generalidad, se define un tensor como una aplicación $p+q$ veces lineal del conjunto $(E)^p \times (E^*)^q$ en K, designando por E^* un dual de E.

Sea f una aplicación [1] de $(E)^p$ en su cuerpo de escalares. Se llama *transformada de f por la sustitución s*, la aplicación que hace corresponder a $(v_1, v_2, \ldots, v_p) \in (E)^p$ el elemento $f(sv_1, sv_2, \ldots, sv_p)$ de K. Si sf designa esta aplicación, se tendrá que

$$sf(v_1, v_2, \ldots, v_p) = f(sv_1, sv_2, \ldots, sv_p).$$

Se establecen inmediatamente las propiedades que siguen:

1.ª Si f es una aplicación p veces lineal, sf también lo es.

2.ª Si s_1 y s_2 son dos sustituciones, se verifica que:

$$s_2(s_1 f) = (s_2 s_1) f.$$

En efecto,

$$\begin{aligned} s_2(s_1 f) &= s_2 f(s_1 v_1, s_1 v_2, \ldots, s_1 v_p) \\ &= f[s_2(s_1 v_1), \ldots, s_2(s_1 v_p)] \\ &= f[(s_2 s_1) v_1, \ldots, (s_2 s_1) v_p] = (s_2 s_1) f. \end{aligned}$$

Aplicación simétrica.—La aplicación f se llama *simétrica* si, cualquiera que sea s, se verifica que

$$sf = f.$$

En la práctica, se asegura la simetría comprobando la relación anterior para todas las trasposiciones de $\mathcal{S}$. El estudio de una aplicación bilineal y simétrica se examinará en el capítulo 17.

4. Aplicación antisimétrica.—Sea f una aplicación de $(E)^p$ en K. Se llama antisimétrica si, cualquiera que sea $s \in \mathcal{S}$, se tiene que

$$sf = \epsilon_s f \qquad [1]$$

o sea,

$$f(sv_1, sv_2, \ldots, sv_p) = \epsilon_s f(v_1, v_2, \ldots, v_p)$$

designando por ϵ_s la signatura de la sustitución: $+1$ para una sustitución par y -1 para una impar. El segundo miembro de la igualdad es el producto de $f \in K$ por el escalar $\epsilon_s \in K$. En particular, para toda trasposición t, $\epsilon_t = -1$ y, por consiguiente,

$$f(tv_1, tv_2, \ldots, tv_p) = -f(v_1, v_2, \ldots, v_p).$$

---

[1] No decimos multilineal porque no es necesario para esta definición, y, también, por ciertas propiedades que veremos en los apartados siguientes.

Esta propiedad de las trasposiciones es característica de las aplicaciones antisimétricas. En efecto, una sustitución es par o impar según la paridad del número de las trasposiciones que permiten realizarla. Cada trasposición cambia el signo de f; por tanto, la ley [1] se cumple para toda sustitución.

Se demuestra inmediatamente que la aplicación nula, $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)=0$ para cualesquiera $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, es antisimétrica. Diremos que esta es la aplicación *degenerada*. No es evidente la existencia de aplicaciones antisimétricas no degeneradas. El caso $n=p$ se examinará en la página 371.

Aplicación alternada.—Una aplicación f de $(E)^p$ en K se llama *alternada* cuando es igual a 0 siempre que coinciden dos vectores. El siguiente teorema compara la alternancia y la antisimetría en el caso de funciones multilineales.

TEOREMA.—*Para que una función p veces lineal sea antisimétrica es necesario y suficiente que sea alternada.*

Sea f una función antisimétrica. En la fórmula [1] se elige para s la trasposición de los elementos $\mathbf{v}_i$ y $\mathbf{v}_j$, de donde $\epsilon_s=-1$, y

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_p)+f(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_p)=0.$$

Como esta relación se cumple cualesquiera que sean $\mathbf{v}_i$ y $\mathbf{v}_j$, supongamos en ella $\mathbf{v}_i=\mathbf{v}_j$, de donde

$$2f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_p)=0$$

y podemos dividir por 2 si K no es de característica 2. Se ve que f es alternada.

Recíprocamente, si f es p veces lineal y alternada, se tiene, para dos vectores cualesquiera:

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_i+\mathbf{v}_j, \ldots, \mathbf{v}_j+\mathbf{v}_i, \ldots, \mathbf{v}_p)=0.$$

Se puede descomponer en una suma de cuatro funciones f, de las que dos son nulas en razón de la alternancia. La relación que queda es válida cualesquiera que sean i y j,

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_p)=-f(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_p).$$

Ya que la antisimetría se verifica en toda trasposición, resulta que f es antisimétrica. En lo que sigue, emplearemos indistintamente las palabras alternada y antisimétrica.

Consecuencia.—*Toda función p veces lineal y antisimétrica de un espacio vectorial toma el valor nulo para todo sistema de p vectores dependientes.*

El vector $\mathbf{v}_p$, p. ej., es combinación lineal de los otros:

$$\mathbf{v}_p = \lambda^1\mathbf{v}_1 + \lambda^2\mathbf{v}_2 + \ldots + \lambda^{p-1}\mathbf{v}_{p-1}.$$

Lo reemplazamos por esta expresión en $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ y utilizamos la linealidad. Los $p-1$ valores de la función f así obtenidos son nulos en virtud de la alternancia.

En particular, en un espacio de n dimensiones, toda función antisimétrica p veces lineal es degenerada si $p > n$.

5. Antisimetrización de una función p veces lineal.—Sea f una función p veces lineal; se llama *antisimetrizada* de f a la función $\mathcal{A}f$ definida por la fórmula

$$\mathcal{A}f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p) = \sum_{s \in \mathcal{S}} \epsilon_s f(s\mathbf{v}_1, s\mathbf{v}_2, \ldots, s\mathbf{v}_p),$$

en la que la sumación del segundo miembro se extiende a las $p!$ sustituciones de $\mathcal{S}$. Como las funciones p veces lineales constituyen un espacio vectorial, esta operación es legítima.

TEOREMA.—*Cualquiera que sea f, la función $\mathcal{A}f$ es antisimétrica.*

Estudiemos el efecto de una sustitución $t \in \mathcal{S}$ sobre $\mathcal{A}f$. Tenemos,

$$\mathcal{A}f(t\mathbf{v}_1, t\mathbf{v}_2, \ldots, t\mathbf{v}_p) = \sum_{s \in \mathcal{S}} \epsilon_s f(ts\mathbf{v}_1, \ldots, ts\mathbf{v}_p)$$

y, cuando s recorre $\mathcal{S}$ por entero, se sabe que $s' = ts$ recorre igualmente todo $\mathcal{S}$. Por otra parte, $\epsilon_{s'} = \epsilon_t\epsilon_s$, o $\epsilon_s = \epsilon_t\epsilon_{s'}$, y, en el segundo miembro, se saca ϵ_t factor común:

$$\mathcal{A}f(t\mathbf{v}_1, t\mathbf{v}_2, \ldots, t\mathbf{v}_p) = \epsilon_t \sum_{s' \in \mathcal{S}} \epsilon_{s'} f(s'\mathbf{v}_1, s'\mathbf{v}_2, \ldots, s'\mathbf{v}_p).$$

La última sumación da de nuevo $\mathcal{A}f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$; por consiguiente,

$$\mathcal{A}f(t\mathbf{v}_1, t\mathbf{v}_2, \ldots, t\mathbf{v}_p) = \epsilon_t \mathcal{A}f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p).$$

La función antisimetrizada es, pues, antisimétrica.

Este teorema permite formar una función antisimétrica a partir de una función dada; recíprocamente, toda función antisimétrica puede obtenerse de esta forma (véase ejercicio 1, a continuación). Sin embargo, el resultado carece de interés salvo si $\mathfrak{H}of$ no es degenerada.

EJERCICIOS:

1. Compárese una función f antisimétrica con su antisimetrizada, y dedúzcase de ello que toda función antisimétrica es la antisimetrizada de otra función.

2. Sea f una función p veces lineal; se define la función simetrizada de f por $Sf = \sum_{s \in \mathfrak{S}_p} f(sv_1, sv_2, \ldots, sv_p)$. Demuéstrese que la simetrizada es simétrica y pruébese también que toda función simétrica puede obtenerse de esta forma.

6. Función determinante en un espacio de n dimensiones.—Nos limitaremos a considerar aquí el caso $p = n$, es decir, las funciones Δ, n veces lineales y antisimétricas, de un espacio de n dimensiones. Tal aplicación se llama muchas veces *determinante de vectores*; sin embargo, emplearemos la expresión *función determinante*, de acuerdo con el vocabulario utilizado en este libro. La existencia de estas funciones, no evidente *a priori*, se deduce de la representación analítica en una base.

Precisemos para el caso presente la fórmula general de la página 366. Tenemos, si Δ es una función de ese tipo, supuesta dada,

$$\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \sum_{\alpha_1, \alpha_2, \ldots, \alpha_n} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} \Delta(\mathbf{e}_{\alpha_1}, \mathbf{e}_{\alpha_2}, \ldots, \mathbf{e}_{\alpha_n}),$$

donde la suma del segundo miembro se extiende a n^n términos. Pero sabemos que $\Delta(\mathbf{e}_{\alpha_1}, \mathbf{e}_{\alpha_2}, \ldots, \mathbf{e}_{\alpha_n})$ es nula si dos de los índices $\alpha_1, \alpha_2, \ldots, \alpha_n$ son iguales (alternancia), por lo que la sumación se reduce a las $n!$ permutaciones de 1, 2, ..., n. Si s es la sustitución $(1, 2, \ldots, n) \longrightarrow (\alpha_1, \alpha_2, \ldots, \alpha_n)$, de donde $\alpha_1 = s1 = s_1$, ... y $\mathbf{e}_{\alpha_1} = s\mathbf{e}_1$ ..., se podrá escribir, en virtud de la antisimetría,

$$\Delta(\mathbf{e}_{\alpha_1}, \mathbf{e}_{\alpha_2}, \ldots, \mathbf{e}_{\alpha_n}) = \Delta(s\mathbf{e}_1, s\mathbf{e}_2, \ldots, s\mathbf{e}_n) = \epsilon_s \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n).$$

La constante $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ se saca factor común, de donde resulta la expresión general,

$$\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) \cdot \sum_{s \in \mathfrak{S}} \epsilon_s x_1^{s_1} x_2^{s_2} \ldots x_n^{s_n}. \qquad [1]$$

Se obtiene así una forma n veces lineal de un tipo muy particular, ya que sus coeficientes son o bien nulos o iguales (salvo el signo) a la constante $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$.

Determinante de una matriz.—Resulta cómodo atribuir un nombre al polinomio que constituye el segundo factor del segundo miembro. Para esto se considera la matriz cuadrada

$$A = [x_i^j]$$

cuyas columnas (índice i) están constituidas por las coordenadas de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ en la base en cuestión. El polinomio

$$\sum_{s \in \mathcal{S}} \epsilon_s x_1^{s_1} x_2^{s_2} \ldots x_n^{s_n}$$

así asociado a la matriz A se llama su *determinante* y se representa por $\Delta(A)$. Se puede, pues, enunciar la siguiente propiedad.

TEOREMA.—*El valor de una función determinante para n vectores referidos a una base, se obtiene multiplicando el determinante de la matriz de las coordenadas de los vectores por el valor de la función para los vectores de la base.*

Recíprocamente, demostremos que la fórmula [1] permite definir todas las funciones determinantes. El estudio es análogo al hecho en la página 347.

TEOREMA DE EXISTENCIA.—*Sea un espacio de n dimensiones referido a una base* $\{\mathbf{e}_i\}$. *La aplicación* Δ, *que hace corresponder a los n vectores* $\mathbf{v}_i = \sum x_i^a \mathbf{e}_a$ *el escalar*

$$\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \Lambda \sum_{s \in \mathcal{S}} x_1^{s_1} x_2^{s_2} \ldots x_n^{s_n} \qquad [1\ bis]$$

donde Λ *es una constante cualquiera, es n veces lineal, antisimétrica y tal que* $\Lambda = \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$. *Es una función determinante.*

La fórmula [1 *bis*] define ahora Δ; su carácter lineal con relación a los n vectores resulta obvio en el segundo miembro. Además, el polinomio bajo el signo $\sum$ representa la antisimetrizada de la aplicación n veces lineal φ que, en la base considerada, asocia a los vectores $\mathbf{v}_i$ (i de 1 a n) el escalar $\varphi(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = x_1^1 x_2^2 \ldots x_n^n$. Se trata, por consiguiente, de una aplicación antisimétrica (véase pág. 370).

Finalmente, si se toma $\mathbf{v}_i = \mathbf{e}_i$ (i de 1 a n), el polinomio contiene un solo término no nulo ($x_{i}^{s_i} = 0$, salvo para $s_i = 1$, etc.), el cual es igual a 1; por tanto, $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) = \Lambda$. Hemos obtenido de este modo una función determinante, que es ciertamente no degenerada si se elige $\Lambda \neq 0$.

Unicidad.—*Existe una función determinante y una sola que toma un valor dado para n vectores linealmente independientes.*

La existencia se deduce del teorema anterior, porque los n vectores $\{\mathbf{e}_i\}$ constituyen una base, y se conoce el valor Λ asociado a esta base. La unicidad resulta de la fórmula [1] que prueba el carácter necesario de la expresión elegida.

7. Consecuencias. 1. Conjunto de las funciones Δ.—Del estudio anterior resulta que se obtiene el conjunto de las funciones determinante atribuyendo a una base dada el conjunto de los posibles valores Λ. Aunque la definición de la función utiliza un sistema de referencia, debemos entenderla asociada intrínsecamente a los vectores considerados. La influencia de un cambio de base se deduce inmediatamente de la fórmula [1], ya que

$$\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) \cdot \Delta(A) = \Delta(\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n) \cdot \Delta(A')$$

designando por $\Delta(A)$ y $\Delta(A')$ los determinantes de las matrices de las coordenadas en la antigua base y en la nueva.

2. Teorema.—*Los valores que toman dos funciones determinante para un mismo sistema arbitrario de n vectores son proporcionales.* Si

$$\Delta_1(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \Delta_1(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) \cdot \Delta(A)$$

y

$$\Delta_2(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \Delta_2(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) \cdot \Delta(A)$$

al hacer el cociente desaparece el determinante de la matriz A y Δ_1/Δ_2 conserva un valor constante.

3. Aplicación degenerada.—TEOREMA.—*Una función determinante es degenerada si, y solo si, es nulo su valor para una base.*

En virtud de la fórmula [1] esta condición es suficiente y es además necesaria, evidentemente.

TEOREMA.—*Para que n vectores de un espacio de n dimensiones*

sean dependientes es necesario y suficiente que anulen a una función determinante Δ *no degenerada de este espacio.*

En efecto; hemos visto que la dependencia implica la nulidad de Δ (pág. 370); recíprocamente, si los vectores son independientes, forman una base y, si es $\Delta = 0$ para esta base, la función es degenerada. Excluida esta hipótesis, será $\Delta \neq 0$ para los n vectores considerados.

8. Orientación de un espacio vectorial.—Suponemos que K es un cuerpo ordenado, aunque en la práctica basta considerar R.

TEOREMA.—*En un espacio de n dimensiones sobre* R, *el conjunto de las bases puede dividirse en dos clases disjuntas y complementarias.*

Elijamos, en efecto, una función Δ no degenerada; por consiguiente, de valor distinto de cero para toda base. Puesto que se trata de números reales, se tendrá o bien $\Delta > 0$ o bien $\Delta < 0$, según la base considerada, y toda base pertenece a una de las dos clases.

Además, la división efectuada es independiente de la elección de Δ, ya que, para pasar de una función a otra, se multiplican los valores por un mismo factor. Las bases que corresponden a $\Delta_1 > 0$ pueden corresponder a $\Delta_2 < 0$, pero la división sigue siendo la misma.

Definición.—Se orienta un espacio vectorial sobre R eligiendo una de las bases anteriores como clase de bases positivas o directas; la otra corresponde a las bases negativas.

En la práctica, para orientar el espacio basta elegir una base considerada como de la clase positiva. Refiriéndonos a la fórmula [1], se ve que para todas las bases directas el determinante de la matriz de las coordenadas (matriz de cambio) con relación a la base elegida es positivo.

EJERCICIOS:

1. En un espacio de n dimensiones, sean v_i y w vectores cualesquiera. Si Δ es una función determinante, pruébese que:

a) $\Delta(v_1, v_2, \ldots, v_n) = \Delta(v_1, v_2 - v_1, \ldots, v_n - v_1)$;
b) $\Delta(v_1, v_1 + \alpha w, v_1 + \beta w, v_4, \ldots, v_n) = 0$;
c) $\Delta[\alpha_1(v_1 - v_2), \alpha_2(v_2 - v_3), \ldots, \alpha_n(v_n - v_1)] = 0$ ($\alpha_1, \alpha_2, \ldots, \alpha_n$, escalares);
d) $\Delta(v_1, v_2, \ldots, v_n) = \Delta(v'_1, v'_2, \ldots, v'_n)$,

con

$$v' = v_i + \alpha_i^1 v_1 + \alpha_i^2 v_2 + \ldots + \alpha_i^{i-1} v_{i-1}.$$

2. Si E es de tres dimensiones, demuéstrese que $2v_2 = v_1 + v_3$ implica

$$\Delta(v_1 + w, v_2 + w, v_3 + w) = 0.$$

3. En el mismo caso, compárense

$$\Delta(v_1, v_2, v_3) \quad y \quad \Delta(v_1 - v_2 - v_3, v_2 - v_1 - v_3, v_3 - v_1 - v_2).$$

DETERMINANTE DE UNA MATRIZ

Según hemos indicado, la definición de determinante de una matriz se ha establecido *a priori*. El estudio que sigue se hace a partir de esta definición y resulta casi independiente de la teoría anterior. Sin embargo, se suponen conocidas las definiciones de las páginas 364 a 367. Para el lector que haya estudiado lo que precede, son innecesarias ciertas demostraciones.

9. Definición.—Sea $A=[x_i^j]$ (i índice de columna) una matriz cuadrada de orden n sobre K. Se llama *determinante* de esta matriz el escalar $\Delta(A)$ definido por la fórmula

$$\Delta(A)=\sum_{s\in S}\epsilon_s x_1^{s_1}x_2^{s_2}\ \dots\ x_n^{s_n}, \qquad [1]$$

donde la suma se extiende a las $n!$ sustituciones de n objetos; se tiene que

$$s=\begin{pmatrix}1 & 2 & \dots & n\\ s_1 & s_2 & \dots & s_n\end{pmatrix}$$

y ϵ_s designa la signatura de esta sustitución.

En el segundo miembro, los números $s_1, s_2, \dots, s_n$ representan una permutación de 1, 2, ..., n; son todos distintos y, por consiguiente, cada término es un producto en el que figura un factor y uno solo de cada columna (subíndice) y de cada fila (superíndice) del cuadro A. Se efectúa la suma de los $n!$ términos así obtenidos, después de multiplicarlos por $+1$ ó -1 según que sea par o impar la sustitución s.

Para los valores $n=2$, 3, la expresión [1] permite fácilmente el cálculo de un determinante. Para $n=2$, se tienen dos sustituciones, de donde

$$\begin{vmatrix} x_1^1 & x_2^1\\ x_1^2 & x_2^2\end{vmatrix} = x_1^1x_2^2 - x_1^2x_2^1.$$

Para $n=3$, hay 6 términos. El lector comprobará la fórmula

$$\begin{vmatrix} x_1^1 & x_2^1 & x_3^1\\ x_1^2 & x_2^2 & x_3^2\\ x_1^3 & x_2^3 & x_3^3\end{vmatrix} = x_1^1x_2^2x_3^3 - x_1^1x_2^3x_3^2 + x_1^2x_2^3x_3^1 - x_1^2x_2^1x_3^3 + x_1^3x_2^1x_3^2 - x_1^3x_2^2x_3^1.$$

Finalmente, cualquiera que sea n, de la fórmula [1] resulta que la matriz unidad de orden n admite un determinante igual a 1.

Notaciones.—Se escribe $\Delta(A)=|x_j^i|$ colocando entre barras (en lugar de corchetes) el cuadro que define la matriz. Por otra parte, se usa la terminología del cálculo matricial: filas, columnas, orden, etc., de un determinante.

Con el lenguaje de la sección precedente, se ve que $\Delta(A)$ es el valor, para los n vectores columna de A, de la función determinante del espacio K^n que toma el valor 1 para la base canónica. Es, pues, no degenerada y, por tanto (pág. 373), una condición necesaria y suficiente para que los n vectores columna de una matriz sean dependientes es que su determinante sea nulo.

10. Determinantes de dos matrices traspuestas. TEOREMA.—*Los determinantes asociados a dos matrices traspuestas son iguales.*

Sean $A=[x_j^i]$ y ${}^tA=[x'^i_j]$, con $x'^i_j=x^j_i$, dos matrices traspuestas. Sus determinantes tienen, respectivamente, por expresiones

$$\Delta=\sum_{s\in\mathcal{S}}\epsilon_s x^1_{s_1}x^2_{s_2}\dots x^n_{s_n} \qquad \text{y} \qquad \Delta'=\sum_{s\in\mathcal{S}}\epsilon_s x^{s_1}_1 x^{s_2}_2\dots x^{s_n}_n,$$

reemplazando en la segunda igualdad las x' en función de las x.

Consideremos un término cualquiera de Δ. Sin alterar su valor podemos modificar el orden de sus factores de manera que los superíndices tomen el orden natural 1, 2, ..., n. Con esto se vuelve a efectuar la sustitución t, inversa de s, en la permutación $s_1, s_2, \dots, s_n$. Los subíndices experimentan la misma sustitución: $1\longrightarrow t1=t_1$, ... Observando que $\epsilon_s=\epsilon_t$ y que, si s recorre $\mathcal{S}$, t lo recorre igualmente, el determinante Δ se escribe:

$$\Delta=\sum_{t\in\mathcal{S}}\epsilon_t x^1_{t_1}x^2_{t_2}\dots x^n_{t_n}=\Delta'$$

Consecuencia.—Vamos a precisar las propiedades de un determinante haciendo que los vectores columna de su matriz desempeñen un papel particular. El teorema anterior permite afirmar que se aplican las mismas propiedades respecto de los vectores fila.

11. Linealidad y antisimetría. TEOREMA 1.—*El determinante de una matriz es una función lineal de cada uno de sus vectores columna.*

Ya se ha indicado que cada término de la fórmula [1] contiene un factor y uno solo de cada columna. Actuando sobre esta columna no se modifican los otros factores, de donde se deduce inmediatamente la linealidad:

a) Si todos los elementos de una columna se multiplican por un mismo escalar λ, el determinante queda multiplicado por λ.

b) Si el i-ésimo vector columna es suma de dos vectores, el determinante es la suma de los determinantes obtenidos tomando por i-ésimo vector columna cada uno de los vectores considerados, sin modificar los demás.

TEOREMA 2.—*La permutación de dos vectores columna transforma un determinante en su opuesto.*

Sean A y A' dos matrices, que se deducen una de otra por permutación de las columnas que ocupan los lugares i y j. Se tiene:

$$\Delta(A) = \sum_{s \in \mathcal{S}} \epsilon_s x_1^{s_1} \dots x_i^{s_i} \dots x_j^{s_j} \dots x_n^{s_n}$$

y

$$\Delta(A') = \sum_{s \in \mathcal{S}} \epsilon_s x_1^{s_1} \dots x_j^{s_i} \dots x_i^{s_j} \dots x_n^{s_n}.$$

Sin alterar el valor de un término de $\Delta(A')$ se pueden cambiar en él los factores $x_j^{s_i}$ y $x_i^{s_j}$, lo que da:

$$\Delta(A') = \sum_{s \in \mathcal{S}} \epsilon_s x_1^{s_1} \dots x_i^{s_j} \dots x_j^{s_i} \dots x_n^{s_n}.$$

Pero ϵ_s no es la signatura de la sustitución s' tal que $s'_1 = s_1, \dots, s'_i = s_j, \dots, s'_j = s_i, \dots, s'_n = s_n$. Evidentemente, $\epsilon_s = -\epsilon_{s'}$ y, por consiguiente,

$$\Delta(A') = -\sum_{s' \in \mathcal{S}} \epsilon_{s'} x_1^{s'_1} \dots x_2^{s'_2} \dots x_n^{s'_n}.$$

Cuando s recorre $\mathcal{S}$, s' recorre igualmente $\mathcal{S}$, y en el segundo miembro se tiene el opuesto del determinante de A, de donde $\Delta(A') = -\Delta(A)$.

Consecuencias.—a) *Un determinante es nulo si dos vectores columna son idénticos o proporcionales.*

En el segundo caso, se aplica primero la linealidad con el fin de

reducirlo a dos vectores idénticos. El determinante debe ser entonces igual a su opuesto.

b) *No se altera el determinante de una matriz si se suma a una columna cualquiera una combinación lineal de las otras columnas.*

El determinante obtenido se descompone en suma de determinantes, utilizando la linealidad. Se obtiene así el determinante inicial y otros determinantes, todos nulos según *a)*.

c) *Un determinante es nulo si sus vectores columna son dependientes* (su matriz es singular).

Una de las columnas es combinación lineal de las otras y el determinante se descompone en suma de determinantes nulos.

12. Determinante de un producto de matrices. TEOREMA.—*El determinante asociado a la matriz producto de dos matrices cuadradas de orden n es igual al producto de los determinantes de estas matrices.*

Sean A y B las matrices dadas. Formemos el producto AB considerando la acción de B sobre las columnas $A_1, A_2, \ldots, A_n$ de A:

$$[A_1\, A_2 \ldots A_n] \begin{bmatrix} b_1^1 & b_2^1 & \ldots & b_n^1 \\ \vdots & \vdots & & \vdots \\ b_1^n & b_2^n & \ldots & b_n^n \end{bmatrix} = [\Sigma b_1^\alpha A_\alpha \;\; \Sigma b_2^\alpha A_\alpha \ldots \Sigma b_n^\alpha A_\alpha].$$

Se debe calcular un determinante cuyas columnas son combinaciones lineales de las de A. La transformación se efectúa como en la página 371. La linealidad conduce a una suma de n^n determinantes, que se escribe:

$$\Delta(AB) = \sum_{\alpha_1, \alpha_2, \ldots, \alpha_n} b_1^{\alpha_1} b_2^{\alpha_2} \ldots b_n^{\alpha_n} \Delta(A_{\alpha_1} A_{\alpha_2} \ldots A_{\alpha_n}).$$

En el segundo miembro se puede reducir primero la sumación a las $n!$ permutaciones de 1, 2, ..., n, porque si dos índices α_i y α_j son iguales, el determinante $\Delta(A_{\alpha_1} A_{\alpha_2} \ldots A_{\alpha_n})$ es nulo. Se restablece a continuación en cada término el orden normal de las columnas de A mediante trasposiciones de columnas. Cada trasposición cambia el signo y, si ϵ_s es la signatura de la sustitución

$$s = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix} \quad \text{(de donde } \alpha_1 = s_1,\ \alpha_2 = s_2, \ldots),$$

se obtiene:

$$\Delta(AB) = \left(\sum_{\lambda \in J} \epsilon_\lambda b^1_{j_1} b^2_{j_2} \dots b^n_{j_n} \right) \Delta(A).$$

Como el primer factor es igual al determinante $\Delta(B)$, se tiene, en efecto,

$$\Delta(AB) = \Delta(A) \cdot \Delta(B).$$

Consecuencia.—*Para que una matriz cuadrada sea regular es necesario y suficiente que su determinante sea distinto de cero.*

Ante todo, si $\Delta(A) \neq 0$ la matriz A es regular, porque hemos visto que toda matriz singular tiene determinante nulo [véase pág. 378, *c)*].

Supongamos ahora que A es regular y demostremos que $\Delta(A) \neq 0$. La matriz A admite una inversa A^{-1}, tal que $AA^{-1} = I$; pasando a determinantes, resulta:

$$\Delta(A) \cdot \Delta(A^{-1}) = \Delta(I) = 1.$$

De donde se deduce que $\Delta(A)$ y $\Delta(A^{-1})$ son distintos de cero e inversos uno del otro. Este teorema proporciona un criterio utilizable en la práctica para estudiar la regularidad de una matriz de orden poco elevado.

OBSERVACIONES.— 1.ª Utilizando el teorema de la página 376 se ve que una cualquiera de las matrices AB, ${}^tAB$, $A'B$, ${}^tA'B$ (o sus traspuestas) tiene por determinante $\Delta(A) \cdot \Delta(B)$. De ello resulta que el producto de dos determinantes se puede escribir de cuatro maneras diferentes.

2.ª Siempre es posible escribir el cuadrado de un determinante en forma de determinante simétrico

$$[\Delta(A)]^2 = \Delta({}^tAA) = \Delta(A'A).$$

Aplicación.—Un determinante puede transformarse mediante combinaciones de filas (o de columnas) que no se reduzcan a las sencillas combinaciones examinadas en las páginas 376 y 377. ¿Cómo se modifica entonces el valor del determinante?

El teorema anterior proporciona la respuesta a esta pregunta. Las transformaciones examinadas equivalen a efectuar una multiplicación a la izquierda (o a la derecha) de la matriz del determinante por una matriz cuadrada del mismo orden: el determinante queda multiplicado por el determinante de esta matriz.

EJEMPLO.—De cada fila de un determinante de tercer orden se resta la suma de las otras dos. Compárense los valores de los determinantes obtenidos.

El determinante inicial se ha multiplicado por $\begin{vmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{vmatrix} = -4.$

EJERCICIOS:

1. Escríbase de cuatro maneras distintas el producto de los dos determinantes

$$\Delta = \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} \quad \text{y} \quad \Delta' = \begin{vmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{vmatrix}.$$

2. Demuéstrese que el determinante

$$\Delta = \begin{vmatrix} a\alpha + b\beta & a'\alpha + b'\beta & a''\alpha + b''\beta \\ a\alpha' + b\beta' & a'\alpha' + b'\beta' & a''\alpha' + b''\beta' \\ a\alpha'' + b\beta'' + c & a'\alpha'' + b'\beta' + c' & a''\alpha'' + b''\beta'' + c'' \end{vmatrix}$$

puede considerarse como el producto de dos determinantes.

13. Desarrollo de un determinante por los elementos de una fila o columna.—Un determinante es una forma lineal con relación a los elementos de cualquier fila o columna. Si ponemos en evidencia los elementos de tal fila o columna (aquí la *i-ésima* columna), se escribirá:

$$\Delta(A) = x_i^1 X_1^i + x_i^2 X_2^i + \ldots + x_i^n X_n^i.$$

Designamos por X_j^i el polinomio homogéneo y de grado $n-1$ obtenido sacando factor común x_i^j de todos los términos que lo contienen; X_j^i se llama *cofactor* o *adjunto* del elemento x_i^j. Puesto que se forma un término de Δ tomando un elemento y uno solo de cada fila y de cada columna de la matriz, se ve que X_j^i no depende ni de los elementos de la columna i, ni de los de la fila j. Vamos a demostrar que estos polinomios están relacionados de manera sencilla con determinantes deducidos de Δ, sus *menores*.

Definición.—Se llama *menor* del elemento x_i^j de Δ el determinante de orden $n-1$ obtenido suprimiendo la columna i y la fila j de Δ sin alterar el orden de las demás.

Cálculo de X_j^i.—Es el coeficiente de x_i^j. Todos los términos que con-

tienen este elemento corresponden a sustituciones s tales que $s1=1$. Sea $\mathcal{S}_1$ el conjunto de estas sustituciones; se tiene:

$$X_1^1 = \sum_{s \in \mathcal{S}_1} \epsilon_s x_2^{s_2} x_3^{s_3} \dots x_n^{s_n}.$$

Para $s \in \mathcal{S}_1$, los números $s_2, s_3, \dots, s_n$ forman una permutación de $2, 3, \dots, n$. Si es σ la sustitución $(2, 3, \dots, n) \longrightarrow (s_2, s_3, \dots, s_n)$ (de donde $s_2=\sigma_2, \dots$), será $\epsilon_s=\epsilon_\sigma$ $(s \in \mathcal{S}_1)$ ya que correspondiéndose el número 1 consigo mismo, las trasposiciones cuyo producto da σ, realizan también s. De ello se deduce que

$$X_1^1 = \sum_{\sigma \in \mathcal{S}_1} \epsilon_\sigma x_2^{\sigma_2} \dots x_n^{\sigma_n}.$$

El segundo miembro es un determinante: el menor de x_1^1.

Cálculo de X_j^i.—Se reduce al caso anterior desarrollando otro determinante Δ', igual u opuesto a Δ y en el que el elemento x_i^j figura en primer lugar. Basta efectuar cambios de filas o de columnas sucesivas: $i-1$ cambios de columnas, seguidos de $j-1$ cambios de filas llevan el elemento considerado al primer lugar, conservando el orden de las restantes filas y columnas. El determinante obtenido es $\Delta' = (-1)^{i+j}\Delta$ y en su desarrollo el coeficiente de x_1^1 es igual a $(-1)^{i+j}X_j^i$. Resulta, pues,

$$(-1)^{i+j}X_j^i = D_j^i,$$

designando por D_j^i al menor de x_i^j escrito en el determinante inicial. Se deduce de ello que

$$X_j^i = (-1)^{i+j}D_j^i.$$

TEOREMA.—*El adjunto del elemento x_i^j es igual u opuesto a su menor según que $i+j$ sea par o impar.*

Un determinante es igual a la suma de los productos de los elementos de una fila o de una columna por sus adjuntos.

14. Aplicaciones. *a)* La regla anterior reduce el cálculo de un determinante de orden n al de n determinantes de orden $n-1$, después al de $n(n-1)$ determinantes de orden $n-2$, etc. Proporciona así un procedimiento sistemático para el cálculo de los determinantes;

no obstante, cuando n es mayor que 3, este procedimiento no es aconsejable. Tenemos, p. ej.,

2.° orden:

$$\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - ba';$$

3.er orden:

$$\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} = a\begin{vmatrix} b' & c' \\ b'' & c'' \end{vmatrix} - b\begin{vmatrix} a' & c' \\ a'' & c'' \end{vmatrix} + c\begin{vmatrix} a' & b' \\ a'' & b'' \end{vmatrix}$$

$$= a(b'c'' - c'b'') - b(a'c'' - c'a'') + c(a'b'' - b'a'').$$

b) Un determinante en el que todos los elementos de una fila o de una columna sean nulos, salvo uno, que es igual a 1, se reduce a un determinante de orden inmediatamente inferior:

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ 0 & 1 & 0 & 0 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} = - \begin{vmatrix} a_1 & a_3 & a_4 \\ b_1 & b_3 & b_4 \\ c_1 & c_3 & c_4 \end{vmatrix}.$$

c) Recíprocamente, esto permite elevar el orden de un determinante. Así, p. ej.,

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \\ \alpha & \beta & 1 \end{vmatrix}.$$

Los elementos α y β pueden ser cualesquiera; en particular, nulos.

d) Un determinante, en el que son nulos todos los elementos situados a un lado de la diagonal principal, es igual al producto de los elementos de esta diagonal principal. Los elementos situados al otro lado no desempeñan papel alguno.

EJERCICIOS:

1. Si $\Delta = |a_j^i|$, pruébese que

$$D = \begin{vmatrix} a_1^1 + x & a_1^2 + x & \dots & a_1^n + x \\ a_2^1 + x & a_2^2 + x & \dots & a_2^n + x \\ \vdots & \vdots & & \vdots \\ a_n^1 + x & a_n^2 + x & \dots & a_n^n + x \end{vmatrix} = \Delta + x \sum_{ij} A_j^i,$$

donde A_j^i es el cofactor (adjunto) de a_j^i en Δ.

2. Demuéstrese la fórmula

$$\begin{vmatrix} a_1+x & x & \dots & x \\ x & a_2+x & \dots & x \\ \vdots & \vdots & & \vdots \\ x & x & \dots & a_n+x \end{vmatrix} = a_1a_2 \dots a_nx \left(\frac{1}{x}+\frac{1}{a_1}+\dots+\frac{1}{a_n} \right),$$

que es un caso particular de la anterior.

3. Calcúlese el determinante

$$\Delta = \begin{vmatrix} a_1^1-\lambda & a_1^2 & a_1^3 \\ a_2^1 & a_2^2-\lambda & a_2^3 \\ a_3^1 & a_3^2 & a_3^3-\lambda \end{vmatrix}$$

y ordénese con relación a λ. Pruébese que los coeficientes de λ dependen de manera sencilla del determinante obtenido del Δ haciendo $\lambda=0$.

4. Demuéstrese que

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \cdot \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1+b_1\beta_1 & a_1\alpha_2+b_1\beta_2 & c_1 & d_1 \\ a_2\alpha_1+b_2\beta_1 & a_2\alpha_2+b_2\beta_2 & c_2 & d_2 \\ a_3\alpha_1+b_3\beta_1 & a_3\alpha_2+b_3\beta_2 & c_3 & d_3 \\ a_4\alpha_1+b_4\beta_1 & a_4\alpha_2+b_4\beta_2 & c_4 & d_4 \end{vmatrix}$$

5. Llamamos Δ' el determinante obtenido reemplazando en el determinante

$$\Delta = \begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ \alpha & \beta & \gamma & \delta \end{vmatrix}$$

los elementos de la cuarta fila por sus cofactores A, B, C, D. Indíquese la expresión de Δ' en función de A, B, C, D. Efectuando el cuadrado de Δ', pruébese que

$$\Delta' = \begin{vmatrix} \Sigma a^2 & \Sigma aa' & \Sigma aa'' \\ \Sigma aa' & \Sigma a'^2 & \Sigma a'a'' \\ \Sigma aa'' & \Sigma a'a'' & \Sigma a''^2 \end{vmatrix},$$

donde $\Sigma a^2 = a^2+b^2+c^2+d^2$, etc.

6. Demuéstrese que

$$D = \begin{vmatrix} \alpha a+\beta a'+\gamma a'' & \alpha b+\beta b'+\gamma b'' \\ \alpha' a+\beta' a'+\gamma' a'' & \alpha' b+\beta' b'+\gamma' b'' \end{vmatrix}$$

puede escribirse en forma de producto de determinantes de tercer orden.

15. Cálculo de un determinante. **Primer método.**—Desarrollo por los elementos de una fila o de una columna. En la práctica, este método no puede utilizarse más que si el orden es pequeño ($n \leqslant 4$) o si

la fila o columna considerada contiene muchos elementos iguales a cero. Sea, p. ej.,

$$\begin{vmatrix} 1+\alpha^2 & \alpha\beta & \alpha\gamma \\ \alpha\beta & 1+\beta^2 & \beta\gamma \\ \alpha\gamma & \beta\gamma & 1+\gamma^2 \end{vmatrix} = (1+\alpha^2)\begin{vmatrix} 1+\beta^2 & \beta\gamma \\ \beta\gamma & 1+\gamma^2 \end{vmatrix} - \alpha\beta\begin{vmatrix} \alpha\beta & \beta\gamma \\ \alpha\gamma & 1+\gamma^2 \end{vmatrix} +$$

$$+\alpha\gamma\begin{vmatrix} \alpha\beta & 1+\beta^2 \\ \alpha\gamma & \beta\gamma \end{vmatrix} = 1+\alpha^2+\beta^2+\gamma^2.$$

Segundo método.—Se transforma la matriz mediante combinaciones de filas o de columnas, teniendo en cuenta cualquier cambio eventual del valor del determinante. El empleo de las operaciones elementales permite anular todos los elementos de una fila o de una columna salvo uno de ellos, lo que nos conduce a utilizar el método anterior. Así, por ejemplo,

$$D = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \\ 1 & 5 & 15 & 35 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 6 \\ 0 & 1 & 4 & 10 \\ 0 & 1 & 5 & 15 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 6 \\ 1 & 4 & 10 \\ 1 & 5 & 15 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 6 \\ 0 & 1 & 4 \\ 0 & 1 & 5 \end{vmatrix} = 1.$$

Si el determinante es un polinomio con relación a ciertas variables, la elección de combinaciones apropiadas pone a veces en evidencia un mismo factor en todos los elementos de una fila o columna. Se puede entonces reducir el orden del determinante en una unidad. Calculemos, p. ej., el determinante

$$\Delta = \begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}.$$

Sumando las dos últimas filas a la primera, resulta:

$$\Delta = \begin{vmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c)\begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

Restemos los elementos de la primera columna de los de las otras:

$$\Delta = (a+b+c)\begin{vmatrix} 1 & 0 & 0 \\ 2b & -(a+b+c) & 0 \\ 2c & 0 & -(c+a+b) \end{vmatrix} = (a+b+c)^3.$$

EJERCICIOS:

1. Calcúlense los siguientes determinantes:

$$\Delta_1=\begin{vmatrix} 1 & 1 & 1 \\ b+c & c+a & a+b \\ bc & ca & ab \end{vmatrix}, \qquad \Delta_2=\begin{vmatrix} x & \frac{l}{m}+\frac{m}{l} & \frac{n}{l}+\frac{l}{n} \\ \frac{l}{m}+\frac{m}{l} & x & \frac{m}{n}+\frac{n}{m} \\ \frac{n}{l}+\frac{l}{n} & \frac{m}{n}+\frac{n}{m} & x \end{vmatrix}.$$

2. Demuéstrese que

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix}=(af-be+cd)^2.$$

3. Calcúlense los determinantes que siguen:

$$\Delta=\begin{vmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{vmatrix}, \qquad \Delta'=\begin{vmatrix} a & b & c & d \\ -1 & x & 0 & 0 \\ 0 & -1 & x & 0 \\ 0 & 0 & -1 & x \end{vmatrix}.$$

4. Calcúlese

$$D=\begin{vmatrix} z & -z & 0 \\ 0 & z^2 & -1 \\ 1 & z & z+1 \end{vmatrix}$$

sabiendo que $z^5=1$.

5. Escríbanse los determinantes

$$\begin{vmatrix} 1 & \cos x & \cos 2x \\ 1 & \cos y & \cos 2y \\ 1 & \cos z & \cos 2z \end{vmatrix}, \qquad \begin{vmatrix} 1 & \cos x & \cos 3x \\ 1 & \cos y & \cos 3y \\ 1 & \cos z & \cos 3z \end{vmatrix}$$

en forma de producto.

6. Demuéstrense las igualdades

$$\begin{vmatrix} (b+c)^2 & b^2 & c^2 \\ a^2 & (c+a)^2 & c^2 \\ a^2 & b^2 & (a+b)^2 \end{vmatrix}=2abc(a+b+c)^3,$$

$$\begin{vmatrix} -a & b & c & d \\ b & -a & d & c \\ c & d & -a & b \\ d & c & b & -a \end{vmatrix}=-(a+b+c-d)(b+c+d-a)(c+d+a-b)(d+a+b-c).$$

$$\begin{vmatrix} x & a & b & x \\ a & x & x & b \\ b & x & x & a \\ x & b & a & x \end{vmatrix} = (b-a)^2 (a+b-2x)(a+b+2x).$$

APLICACION A LOS SISTEMAS LINEALES

16. Identidades preliminares.—Sea $\Delta(A) = |x_i^j|$ el determinante de una matriz de orden n. A cada término x_i^j asociamos su cofactor X_j^i. En virtud del teorema de la página 381, se puede escribir:

$$\Delta(A) = \sum_{\alpha=1}^{\alpha=n} x_i^\alpha X_\alpha^i$$

y

$$\Delta(A) = \sum_{\alpha=1}^{\alpha=n} x_\alpha^i X_i^\alpha \qquad (i \text{ entero de } 1 \text{ a } n).$$

Consideremos ahora las sumas análogas a las precedentes, pero que dependen de dos enteros i y j supuestos distintos:

$$\Delta_1 = \sum_{\alpha=1}^{\alpha=n} x_i^\alpha X_\alpha^j \qquad \text{y} \qquad \Delta_2 = \sum_{\alpha=1}^{\alpha=n} x_\alpha^i X_j^\alpha.$$

En la primera, los cofactores son los de la columna j de A, y los términos, los de la columna i; Δ_1 es, por consiguiente, el desarrollo de un determinante cuya matriz se deduce de A reemplazando la columna j por la *i-ésima* sin cambiar las restantes columnas que han servido para formar los cofactores. De ello resulta que la matriz contiene dos columnas idénticas, por lo que su determinante es nulo.

Asimismo, la segunda suma, Δ_2, es nula porque es el desarrollo de un determinante que contiene dos veces una misma fila (la *i-ésima* de A). Utilizando el símbolo de Kronecker, δ_i^j, los resultados anteriores se condensan en las fórmulas:

$$\sum_{\alpha=1}^{\alpha=n} x_i^\alpha X_\alpha^j = \delta_i^j \Delta \qquad \text{y} \qquad \sum_{\alpha=1}^{\alpha=n} x_\alpha^i X_j^\alpha = \delta_j^i \Delta.$$

17. Inversión de una matriz.—El empleo de los determinantes permite dar forma explícita al término general de la inversa de una matriz regular $A = [x_i^j]$.

Matriz adjunta.—Se denomina *adjunta* de una matriz cuadrada $A=[x_i^j]$, la matriz A^+ en la que el término de la columna i y de la fila j es el cofactor del elemento de la fila i y de la columna j de A. Esta es la traspuesta de la matriz obtenida reemplazando en A cada término por su cofactor; el término homólogo de x_j^i en A es X_j^i en A^+. Estudiemos la matriz producto AA^+. Su término de la fila i y columna j es, de acuerdo con las fórmulas precedentes,

$$\sum_{\alpha=1}^{\alpha=n} x_\alpha^i X_j^\alpha = \delta_j^i \Delta.$$

La matriz AA^+ es, pues, escalar, siendo los términos de la diagonal iguales a $\Delta(A)$. Se obtiene el mismo resultado para A^+A:

$$AA^+ = A^+A = (\Delta)I.$$

Inversa.—Se obtiene A^{-1} dividiendo los dos miembros de la igualdad anterior por $\Delta(A) \neq 0$, lo que da:

$$A^{-1} = \frac{A^+}{\Delta(A)}$$

Se deduce así la siguiente regla: el término general (fila i, columna j) de la inversa de una matriz A es igual al cofactor del elemento situado en la fila j y columna i de A, dividido por el determinante de la matriz.

EJERCICIOS:

1. Determínense las matrices adjunta e inversa de cada una de las siguientes matrices:

$$\begin{bmatrix} 3 & -1 & 2 \\ 1 & 0 & 3 \\ 4 & 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 & -1 \\ 0 & 1 & -1 \\ 2 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & a & b & c \\ a & 0 & c & b \\ b & c & 0 & a \\ c & b & a & 0 \end{bmatrix}.$$

2. Resuélvanse los ejercicios 1 y 4 (págs. 327 y 328) utilizando el método precedente para formar la inversa. Compárense ambos métodos.

3. Lo mismo para los ejercicios 5, 6 y 7 (pág. 328).

4. Se considera sobre R el conjunto de las matrices cuadradas de orden 2, $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tales que $A^3=I$. Demuéstrese que son invertibles. ¿Cuál es el valor de $|A|$? Escribiendo la igualdad anterior en la forma $A^2=A^{-1}$, caracterícese el conjunto de las matrices estudiadas.

18. Sistema de Cramer.—Se ha demostrado (pág. 304) que un sistema lineal de n ecuaciones con n incógnitas cuya matriz sea regular (sistema de Cramer) admite solución única. Indiquemos una forma elegante de expresar esta solución. Adoptando la notación matricial, el sistema

$$\sum_{\alpha=1}^{\alpha=n} a^i_\alpha x^\alpha = b^i \qquad \text{se escribe} \qquad \mathrm{AX} = \mathrm{B}.$$

Multiplicando a la izquierda por A^{-1}, se obtiene:

$$\mathrm{X} = \mathrm{A}^{-1}\mathrm{B},$$

de donde, según la fórmula anterior, $\mathrm{X} = \dfrac{\mathrm{A}^{+}\mathrm{B}}{\Delta}$. Para obtener la *i-ésima* incógnita x^i se considera la *i-ésima* fila del segundo miembro:

$$x^i = \frac{1}{\Delta} \sum \mathrm{X}^i_\alpha b^\alpha.$$

En el numerador figura el desarrollo del determinante que se deduce de Δ reemplazando en él la *i-ésima* columna por B.

TEOREMA.—*Un sistema de Cramer admite solución única; la incógnita x^i es igual al cociente de dos determinantes: el denominador es el determinante de la matriz* A *de los coeficientes, y el numerador, el determinante de la matriz deducida de* A *reemplazando el i-ésimo vector columna por el vector de los segundos miembros.*

Las fórmulas anteriores resultan poco eficaces para el cálculo numérico a partir de $n \geqslant 4$. En estos casos es preferible utilizar el método de la página 144.

19. Ejemplo.—Resuélvase el sistema

$$\begin{cases} x + y + z = 1, \\ ax + by + bz = m, \\ a^2x + b^2y + c^2z = m^2, \end{cases}$$

en el que suponemos $a \neq b \neq c$.

El determinante de las incógnitas es

$$\mathrm{A} = \begin{vmatrix} 1 & 1 & 1 \\ a & b & b \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)\,(c^2-b^2).$$

Por consiguiente, es diferente de cero; se obtiene:

$$x=\frac{1}{\Delta}\begin{vmatrix}1 & 1 & 1\\ m & b & b\\ m^2 & b^2 & c^2\end{vmatrix}=\frac{(c^2-b^2)(b-m)}{(c^2-b^2)(b-a)}=\frac{b-m}{b-a},$$

$$y=\frac{1}{\Delta}\begin{vmatrix}1 & 1 & 1\\ a & m & b\\ a^2 & m^2 & c^2\end{vmatrix}=\frac{(c^2-ab+ma-mb)(m-a)}{(c^2-b^2)(b-a)},$$

$$z=\frac{1}{\Delta}\begin{vmatrix}1 & 1 & 1\\ a & b & m\\ a^2 & b^2 & m^2\end{vmatrix}=\frac{(m-a)(m-b)}{(c^2-b^2)}$$

EJERCICIOS:

1. Resuélvanse los sistemas siguientes:

$$\begin{cases}x+ay-a^2z=a^4,\\ x+by-b^2z=b^4,\\ x+cy-c^2z=c^4.\end{cases}\qquad\begin{cases}\dfrac{x}{h-a}+\dfrac{y}{h-b}+\dfrac{z}{h-c}=1\\ \dfrac{x}{h'-a}+\dfrac{y}{h'-b}+\dfrac{z}{h'-c}=1\\ \dfrac{x}{h''-a}+\dfrac{y}{h''-b}+\dfrac{z}{h''-c}=1.\end{cases}$$

2. Fórmese un polinomio de tercer grado que tome los valores 4, 3, 1, 2 para $x=0, 1, 2, 3$.

3. Si Δ y Δ^+ son los determinantes de una matriz A y de su adjunta A^+, demuéstrese que $\Delta^+=\Delta^{n-1}$, donde n designa el orden.

20. Discusión de un sistema de m ecuaciones con n incógnitas.— La teoría del rango de una matriz ha permitido discutir un sistema cualquiera (véase pág. 304). La práctica se reduce a hallar matrices cuadradas regulares, lo que puede hacerse con ayuda de los determinantes. Llamamos determinante principal al determinante de la matriz principal del sistema y determinante ampliado u orlado (se dice también determinante característico) al determinante de una matriz orlada de la matriz principal.

EJEMPLO.—Tres ecuaciones con cuatro incógnitas:

$$\begin{cases}ax+by+cz+dt=e,\\ a'x+b'y+c'z+d't=e',\\ a''x+b''y+c''z+d''t=e''.\end{cases}$$

Discutamos el rango de la matriz de los coeficientes.

1.º *Rango* 3.—Existe al menos un determinante de tercer orden no nulo; p. ej.,

$$\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} \neq 0.$$

No hay determinante ampliado. Se da un valor arbitrario a t, incógnita no principal, y se resuelve el sistema de Cramer en x, y, z. Se pueden expresar también x, y, z en función de t.

2.º *Rango* 2.—Supongamos, p. ej., $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \neq 0$, siendo nulos los cuatro determinantes de tercer orden. El determinante ampliado debe ser nulo:

$$\begin{vmatrix} a & b & f \\ a' & b' & f' \\ a'' & b'' & f'' \end{vmatrix} = 0.$$

Se dan valores arbitrarios a z y t y se resuelve el sistema de las dos primeras ecuaciones con las incógnitas x e y. La tercera ecuación queda automáticamente satisfecha.

3.º *Rango* 1.—Un elemento es no nulo, p. ej., $a \neq 0$. Todos los determinantes de orden 2 son nulos (lo que da seis condiciones). Las condiciones de posibilidad son dos:

$$\begin{vmatrix} a & f \\ a' & f' \end{vmatrix} = 0, \qquad \begin{vmatrix} a & f \\ a'' & f'' \end{vmatrix} = 0.$$

El sistema se reduce a la primera ecuación, en la que y, z, t se eligen arbitrariamente.

EJERCICIOS:

1. Resuélvanse y discútanse los sistemas

$$\begin{cases} 2x-y+3z=1, \\ -4x+2y+z=3, \\ -2x+y+4z=4, \\ 10x-5y-6z=-10; \end{cases} \qquad \begin{cases} 2x-3y+4z-t=3, \\ x+2y-z+2t=1, \\ 3x-y+2z-3t=4, \\ 3x-y+z-7t=4. \end{cases}$$

2. Demuéstrese que el sistema

$$\begin{cases} x=by+cz+dt, \\ y=cz+dt+ax, \\ z=dt+ax+by, \\ t=ax+by+cz \end{cases}$$

es compatible si, y solo si, se verifica la condición

$$\frac{a}{a+1}+\frac{b}{b+1}+\frac{c}{c+1}+\frac{d}{d+1}=0.$$

3. Resuélvanse y discútanse los siguientes sistemas:

$$\begin{cases} \lambda x+ y+ z=1 \\ x+\lambda y+ z=1 \\ x+ y+\lambda z=1, \end{cases} \quad \begin{cases} 2x+ y- z=a \\ x+my+ z=b \\ 3x+ y-mz=c, \end{cases} \quad \begin{cases} bcx+y+bz=\alpha \\ cax+y+az=\beta \\ abx+y+cz=\gamma, \end{cases}$$

$$\frac{ax+by+h}{c}=\frac{bx+cy+h}{a}=\frac{cx+ay+h}{b}$$

4. Estúdiese la independencia de las cuatro formas lineales

$$P=(1-a)x+y+z+t, \quad Q=x+(1-b)y+z+t,$$
$$R=x+y+(1-c)z+t, \quad S=x+y+z+(1-d)t,$$

según los valores de a, b, c, d.

PRODUCTO EXTERIOR

21. Definición.—Sea E un espacio de n dimensiones en el que se ha elegido una función determinante Δ. Se sabe, pues, asociar a todo sistema ordenado $(v_1, v_2, \ldots, v_n)$ un escalar $\Delta(v_1, v_2, \ldots, v_n)$. Supongamos que $v_1, v_2, \ldots, v_{n-1}$ permanecen fijos y que $v_n = w$ describe E. Las propiedades de Δ prueban que la función así obtenida

$$f(w)=\Delta(v_1, v_2, \ldots, v_{n-1}, w)$$

es lineal en w. Es una aplicación lineal de E en su cuerpo base. Si E^* designa un espacio dual de E, existe un vector $w^* \in E^*$ bien determinado tal que, para w arbitrario, se verifique

$$f(w)=\langle w, w^*\rangle.$$

Se asocia el vector w^* así obtenido a todo sistema de $n-1$ vectores $\{v_1, \ldots, v_{n-1}\}$ de E. Hemos definido, pues, una aplicación de $(E)^{n-1}$ en E^*. El vector w^* se llama *producto exterior* de los vectores $v_1, v_2, \ldots, v_{n-1}$, y se escribe:

$$w^*=[v_1, v_2, \ldots, v_{n-1}]$$

y se tiene, para cualquier w en E,

$$\langle w, [v_1, v_2, \ldots, v_{n-1}]\rangle = \Delta(v_1, \ldots, v_{n-1}, w). \qquad [1]$$

La definición precedente supone la elección de una función determinante. Si Δ se reemplaza por $\lambda\Delta$, $\mathbf{w}^*$ queda reemplazado por $\lambda\mathbf{w}^*$.

22. Propiedades.—1. *La aplicación de* $(E)^{n-1}$ *en* E^* *definida por*

$$\mathbf{w}^* = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}]$$

es $n-1$ *veces lineal.*

Demostremos la linealidad con relación a $\mathbf{v}_1$. Si a $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ corresponde $\mathbf{w}_1^*$ y a $\mathbf{v}'_1, \ldots, \mathbf{v}_{n-1}$ corresponde $\mathbf{w}_2^*$, a $\lambda\mathbf{v}_1 + \mu\mathbf{v}'_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$ corresponderá $\mathbf{w}^*$ tal que, para $\mathbf{w}$ arbitrario,

$$\begin{aligned}\langle \mathbf{w}, \mathbf{w}^* \rangle &= \Delta(\lambda\mathbf{v}_1 + \mu\mathbf{v}'_1, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) \\ &= \lambda\Delta(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) + \mu\Delta(\mathbf{v}'_1, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) \\ &= \lambda\langle \mathbf{w}, \mathbf{w}_1^* \rangle + \mu\langle \mathbf{w}, \mathbf{w}_2^* \rangle = \langle \mathbf{w}, \lambda\mathbf{w}_1^* + \mu\mathbf{w}_2^* \rangle\end{aligned}$$

y, por consiguiente, puesto que $\mathbf{w}$ es arbitrario, $\mathbf{w}^* = \lambda\mathbf{w}_1^* + \mu\mathbf{w}_2^*$.

2. *La aplicación de* $(E)^{n-1}$ *en* E^* *definida por* $\mathbf{w}^* = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}]$ *es alternada.*

Si dos vectores $\mathbf{v}_i$ y $\mathbf{v}_j$ (i y j de 1 a $n-1$) son iguales,

$$\langle \mathbf{w}, \mathbf{w}^* \rangle = \Delta(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) = 0$$

y, puesto que $\mathbf{w}$ es arbitrario, ello implica $\mathbf{w}^* = \mathbf{0}$.

3. *Una condición necesaria y suficiente para que el producto exterior de* $n-1$ *vectores sea nulo es que estos vectores sean dependientes.*

Hemos supuesto la función determinante Δ no degenerada; por consiguiente, la relación $\Delta(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) = 0$ exige que los n vectores sean dependientes. Pero, puesto que estamos en un espacio de n dimensiones y $\mathbf{w}$ es arbitrario, para ello es necesario y suficiente que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ sean dependientes.

4. *El producto exterior de* $n-1$ *vectores es ortogonal a cada uno de estos vectores.*

La relación $\langle \mathbf{v}_i, \mathbf{w}^* \rangle = 0$ para i de 1 a $n-1$, resulta inmediatamente de la fórmula [1].

23. Determinación en una base.—Supongamos E y E^* referidos a dos bases duales $\{\mathbf{e}_i\}$ y $\{\mathbf{e}^{*i}\}$. Se eligen $n-1$ vectores de E:

$$\mathbf{v}_i = \sum x_i^\alpha \mathbf{e}_\alpha \qquad (i \text{ de } 1 \text{ a } n-1)$$

y se pide determinar las coordenadas en la base dual del producto exterior $\mathbf{w}^* = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}]$.

Haciendo

$$\mathbf{w}^* = \sum \xi_\alpha \mathbf{e}^{*\alpha} \quad \text{y} \quad \mathbf{w} = \sum x^\alpha \mathbf{e}_\alpha,$$

la relación [1] se escribe:

$$\sum_1^n x^\alpha \xi_\alpha = \Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) = \begin{vmatrix} x_1^1 & x_2^1 & \ldots & x_{n-1}^1 & x^1 \\ x_1^2 & x_2^2 & \ldots & x_{n-1}^2 & x^2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^n & x_2^n & \ldots & x_{n-1}^n & x^n \end{vmatrix} \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n).$$

Salvo un factor constante $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, el primer miembro es igual al determinante de las coordenadas y representa el desarrollo de este determinante por los elementos de su última columna. Los ξ_α se deducen de los cofactores de los elementos de esta columna multiplicándolos por $\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$.

TEOREMA.—*Las coordenadas del producto exterior de $n-1$ vectores, en la base dual de una base de* E, *son proporcionales a los cofactores de los elementos de la n-ésima columna de un determinante que tiene por $n-1$ primeras columnas las coordenadas de los vectores dados en la base considerada.*

El lector habrá reconocido que esto constituye la generalización, al espacio de n dimensiones, de la teoría del producto vectorial de la geometría de tres dimensiones.

EJERCICIOS:

1. Sean $\{\mathbf{e}_i\}$ y $\{\mathbf{e}^{*i}\}$ dos bases duales de E y E*. Estúdiese el producto exterior de $n-1$ vectores de la base de E. Se demostrará que

$$[\mathbf{e}_1, \ldots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \ldots, \mathbf{e}_n] = (-1)^{n-i} \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) \mathbf{e}^{*i}.$$

2. Sean E_n un espacio de n dimensiones y E_m un espacio de m dimensiones. Una aplicación n veces lineal de E_n en E_m, $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, se supone antisimétrica. Demuéstrese que se puede escribir en la forma:

$$f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \mathbf{w}\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n),$$

en la que $\mathbf{w}$ es un vector determinado de E_m y $\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ una función determinante de los n vectores.

EJERCICIOS SOBRE EL CAPITULO XVI

16-1. Resuélvase la ecuación

$$\begin{vmatrix} a-x & b-x & c-x \\ a'-x & b'-x & c'-x \\ a''-x & b''-x & c''-x \end{vmatrix} = 0.$$

16-2. Demuéstrese que el determinante

$$\begin{vmatrix} 2 & 9 & 9 \\ 4 & 6 & 8 \\ 7 & 4 & 1 \end{vmatrix}$$

es divisible por 13, observando que los números 299, 468, 741 lo son.

16-3. Calcúlense los determinantes

$$D = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}, \qquad \Delta = \begin{vmatrix} C_0^0 & C_1^0 & C_2^0 & \dots & C_{n-1}^0 \\ C_1^1 & C_2^1 & C_3^1 & \dots & C_n^1 \\ C_2^2 & C_3^2 & C_4^2 & \dots & C_{n+1}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ C_{n-1}^{n-1} & C_n^{n-1} & C_{n+1}^{n-1} & \dots & C_{2n-2}^{n-1} \end{vmatrix}$$

de los que el primero es un caso particular del segundo; C_n^p es el número de combinaciones de n objetos tomados p a p. (*Resp.:* $D=\Delta=1$.)

16-4. Se considera el determinante R_n de orden n tal que $a_{ij}=1$ para $i \neq j$ y $a_{ii}=0$ y el determinante D_n deducido del R_n reemplazando por 1 el primer elemento de la diagonal principal.

a) Exprésense D_n y R_n en función de D_{n-1} y R_{n-1}.

b) Dedúzcanse de ello los valores de D_n y de R_n.

[*Resp.:* $D_n=(-1)^{n-1}$; $\quad R_n=(-1)^{n-1}(n-1)$.]

16-5. El término general a_{ij} de un determinante verifica la relación

$$a_{ij}=a_{i-1,j}+a_{i,j-1},$$

que permite calcular todos los elementos cuando se conocen los de la primera fila y primera columna. Se supone, pues, $a_{1,j}=a_{j,1}=j+1$.

Escríbase el determinante Δ_n de orden n que cumple estas condiciones y demuéstrese que verifica la relación de recurrencia

$$\Delta_n=2\Delta_{n-1}-\Delta_{n-2}.$$

Dedúzcase de ella el valor de Δ_n.

16-6. Se considera un determinante de orden n en el que la fila que ocupa el lugar p tiene sus $p-2$ primeros elementos nulos (se supone $p \geqslant 3$). Determí-

nese el número N_n de términos distintos de cero en el desarrollo del determinante.

(*Resp.:* $N_n=2^{n-1}$.)

La misma cuestión suponiendo además que los dos últimos elementos de cada una de las dos primeras filas son nulos.

(*Resp.:* $N'_n=2^{n-1}-4$.)

16-7. Sea un determinante Δ, en el que el término a_{ij} es un polinomio con relación a x; el propio determinante es un polinomio $\Delta(x)$. Hállese su derivada aplicando la definición del número 7 (pág. 118). Demuéstrese que

$$\Delta'=\sum_{i,j} a'_{ij}A_{ij} \quad (i \text{ y } j \text{ varían separadamente de 1 a } n),$$

designando por a'_{ij} la derivada del término a_{ij} y por A_{ij} su cofactor. Demuéstrese que Δ' representa la suma de n determinantes deducidos fácilmente de Δ.

16-8. *Aplicación:* Demuéstrese que el determinante de orden n

$$\Delta_n=\begin{vmatrix} 1+x & 1 & 1 & \dots & 1 \\ 1 & 1+x & 1 & \dots & 1 \\ 1 & 1 & 1+x & \dots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \dots & 1+x \end{vmatrix}$$

satisface la relación $\Delta'_n=n\Delta_{n-1}$. Dedúzcase de ello el valor de Δ_n, hallando primero Δ_2, Δ_3, Δ_4 y observando que $\Delta_n(0)=0$.

16-9. Escríbase el determinante

$$\begin{vmatrix} x^2(y-z)^2 & x^n & 1 \\ y^2(z-x)^2 & y^n & 1 \\ z^2(x-y)^2 & z^n & 1 \end{vmatrix}$$

en forma de producto de cuatro factores, en el que tres son de primer grado en x, y, z, y el cuarto es un polinomio simétrico de grado $n+1$ en x, y, z. Indíquese de manera precisa cómo se forma este último polinomio.

(*Resp.:* Es igual a la suma de todos los polinomios simétricos de la forma $\Sigma x^\alpha y^\beta$, con $\alpha+\beta=n+1$.)

16-10. Sean (x_i, y_i), $i=1, 2, 3$, las coordenadas de un punto M_i en un sistema de coordenadas rectangulares del plano. Demuéstrese que el área algebraica del triángulo $M_1M_2M_3$ es

$$\mathcal{A}=\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

16-11. El mismo problema considerando cuatro puntos $M_i(x_i, y_i, z_i)$, $i=1, 2, 3, 4$, de un espacio de tres dimensiones. Pruébese que el volumen del tetraedro

orientado $M_1M_2M_3M_4$ es

$$V=\frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}.$$

16-12. Sea un tetraedro OABC referido a un sistema de ejes trirrectangulares $Oxyz$, y sean $x_i=\rho_i\cos\alpha_i$, $y_i=\rho_i\cos\beta_i$, $z_i=\rho_i\cos\gamma_i$ $(i=1, 2, 3)$ las coordenadas de los puntos A, B, C. Demuéstrese que el volumen del tetraedro es

$$V=\frac{1}{6}\rho_1\rho_2\rho_3\times\Delta,$$

donde

$$\Delta=\begin{vmatrix} \cos\alpha_1 & \cos\beta_1 & \cos\gamma_1 \\ \cos\alpha_2 & \cos\beta_2 & \cos\gamma_2 \\ \cos\alpha_3 & \cos\beta_3 & \cos\gamma_3 \end{vmatrix}.$$

1.º Demuéstrese que $|\Delta|\leqslant 1$ (el signo $|\ |$ indica valor absoluto). Se podrá razonar bien algebraicamente o geométricamente para hallar el tetraedro de volumen máximo que corresponde a los valores ρ_1, ρ_2, ρ_3 dados.

2.º Dedúzcase de ello la desigualdad

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}^2 \leqslant (x_1^2+y_1^2+z_1^2)(x_2^2+y_2^2+z_2^2)(x_3^2+y_3^2+z_3^2).$$

Desigualdad de M. HADAMARD, válida para un determinante de orden cualquiera.

16-13. 1.º Dados cuatro números complejos

$$\alpha=a+ia', \qquad \beta=b+ib', \qquad \gamma=c+ic', \qquad \delta=d+id'$$

designamos por $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$, $\bar{\delta}$ los números complejos conjugados de α, β, γ, δ, respectivamente.

Dése, bajo forma de polinomio en a, a', b, b', c, c', d, d', el valor del número complejo $D=(\alpha\beta-\gamma\delta)(\bar{\alpha}\bar{\beta}-\bar{\gamma}\bar{\delta})$ efectuando el menor cálculo posible.

2.º Compruébese que D es igual al valor del determinante

$$\Delta=\begin{vmatrix} a & -a' & b & -b' \\ a' & a & b' & b \\ c & -c' & d & -d' \\ c' & c & d' & d \end{vmatrix}.$$

3.º Demuéstrese que D es la suma de los cuadrados de dos polinomios de segundo grado en a, a', b, b', c, c', d, d'. ¿Puede ser el determinante Δ

de rango 3? Resuélvase y discútase el sistema lineal

$$\begin{aligned} ax - a'y + bz - b't &= 0, \\ a'x + ay + b'z + bt &= 0, \\ cx - c'y + dz - d't &= 0, \\ c'x + cy + d'z + dt &= 0, \end{aligned}$$

con coeficientes reales.

4.° Demuéstrese que las matrices con coeficientes reales, de la forma

$$\begin{bmatrix} a & -a' & b & -b' \\ a' & a & b' & b \\ c & -c' & d & -d' \\ c' & c & d' & d \end{bmatrix},$$

constituyen un anillo no conmutativo si se toman como operaciones la adición y la multiplicación de matrices, de las que se supondrán conocidas sus propiedades.

5.° Pruébese que estas matrices forman un espacio vectorial sobre el cuerpo de los números reales, teniendo la multiplicación por un escalar su significado ordinario. Precísese una base de este espacio e indíquese su dimensión.

(Certificat de M.G. . Besançon.)

16-14. 1.° Se considera la ecuación [1] $X^n - 1 = 0$, en la que n es entero positivo. ¿Cuáles son los módulos y los argumentos de las raíces de la ecuación [1]?

2.° En lo sucesivo se designa por ϵ el número complejo $\cos \frac{2\pi}{n} + i \operatorname{sen} \frac{2\pi}{n}$. Si es q un entero, calcúlese la suma $\sum_{j=0}^{n-1} \epsilon^{qj}$, distinguiendo en la misma dos casos según que q sea o no divisible por n.

3.° Se consideran las dos matrices

$$A = [a_{p,q}] \qquad y \qquad B = [b_{p,q}]$$

de n filas y n columnas. Los elementos $a_{p,q}$ y $b_{p,q}$, que figuran en la intersección de la fila de lugar p con la columna de lugar q en las matrices A y B, respectivamente, vienen dados por las fórmulas

$$a_{p,q} = \epsilon^{(p-1)(q-1)} \qquad y \qquad b_{p,q} = \epsilon^{-(p-1)(q-1)}.$$

Calcúlense las matrices productos A^2, B^2, AB y BA.

Calcúlese el producto de los determinantes de A y de B, así como los cuadrados de los determinantes de las matrices A y B.

4.° Calcúlese el producto de los cuadrados de las diferencias dos a dos de las raíces de la ecuación [1]; se precisará el signo de este producto según los valores del resto de la división de n por 4.

5.° Se designa por M la matriz siguiente:

$$\begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_1 & a_2 & \dots & a_0 \\ a_2 & a_3 & \dots & a_1 \\ \dots & \dots & \dots & \dots \\ a_{n-1} & a_0 & \dots & a_{n-2} \end{bmatrix}$$

donde $a_0, \dots, a_{n-1}$ representan n números complejos y en la que cada fila se deduce de la anterior por permutación circular.

Calcúlense las matrices AM, MA, BM, MB, AMA, BMB. Demuéstrese que el determinante de M, considerado como un polinomio de $a_0, \dots, a_{n-1}$, es igual a un producto de n factores de primer grado.

Escríbanse explícitamente estos factores para $n=2$, $n=3$, $n=4$.

(Certificat de M.G.P., Poitiers.)

16-15. Determínese la inversa de la matriz

$$A = \begin{bmatrix} 2 & 0 & 7 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 0 & 4 & 0 \\ 0 & 7 & 0 & 0 \end{bmatrix}.$$

En general, considérese una matriz «en tablero de damas», en la que los cuadros blancos son nulos y los negros no nulos (o al contrario). ¿Es también su inversa una matriz de este tipo?

CAPITULO XVII

OPERADORES Y MATRICES

REPRESENTACION MATRICIAL DE LAS APLICACIONES LINEALES

1. Operadores y matrices.—El estudio hecho en el capítulo 15 de las aplicaciones lineales tiene carácter intrínseco. Vamos a precisarlo utilizando la referencia a las bases. Veremos de esta forma aparecer una importante ligazón entre la teoría de las matrices y la de los operadores.

Sean E y F dos espacios vectoriales de n y m dimensiones, respectivamente; $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ una base de E, y $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ una base de F.

Hemos visto (pág. 337) que una aplicación lineal T de E en F queda completamente determinada eligiendo (arbitrariamente) los transformados de los n vectores de una base de E. Si $\epsilon_1 = T\mathbf{e}_1$, $\epsilon_2 = T\mathbf{e}_2$, ..., $\epsilon_n = T\mathbf{e}_n$, el transformado de $\mathbf{v} = \sum x^\alpha \mathbf{e}_\alpha$ es $\mathbf{w} = T\mathbf{v} = \sum x^\alpha \epsilon_\alpha$.

Utilicemos ahora la referencia de F; se puede escribir:

$$\epsilon_i = \sum_1^m a_i^\alpha \mathbf{f}_\alpha \qquad (i \text{ de } 1 \text{ a } n)$$

y, finalmente, la aplicación T queda definida por los mn coeficientes a_i^j. Resulta cómodo ordenarlos en una matriz de m filas y n columnas: $\mathbf{A} = [a_i^j]$ (i columna, j fila). La columna de lugar i está así constituida por las coordenadas del vector ϵ_i. Se dice que A es la matriz de la aplicación referida a las bases elegidas; quede bien claro que depende de la elección de las bases. Reciprocamente, toda matriz (m, n) define una aplicación de E en F referidos a bases determinadas.

TEOREMA.—*La elección de una base en cada uno de los espacios* E *y* F *establece una biyección entre el conjunto de las aplicaciones lineales de* E *en* F *y el conjunto de las matrices* (m, n). *Las columnas de una matriz son las coordenadas en la base de* F *de los transformados de los vectores de la base de* E.

Escritura matricial.—Estudiemos ahora el transformado **w** de un vector **v** cualquiera. Utilizando las expresiones de los ϵ_i, tenemos:

$$\mathbf{w} = \sum_{1}^{n} x^\beta \epsilon_\beta = \sum_{1}^{n} x^\beta \left(\sum_{1}^{m} a_\beta^\alpha \mathbf{f}_\alpha \right) = \sum_{1}^{m} \sum_{1}^{n} a_\beta^\alpha x^\beta \mathbf{f}_\alpha.$$

Si ponemos $\mathbf{w} = \sum_{1}^{m} y^\alpha \mathbf{f}_\alpha$, se obtienen las fórmulas de transformación de las coordenadas, $y^\alpha = \sum_{1}^{n} a_\beta^\alpha x^\beta$; desarrolladas, se escriben:

$$\begin{cases} y^1 = a_1^1 x^1 + a_2^1 x^2 + \ldots + a_n^1 x^n, \\ y^2 = a_1^2 x^1 + a_2^2 x^2 + \ldots + a_n^2 x^n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y^m = a_1^m x^1 + a_2^m x^2 + \ldots + a_n^m x^n. \end{cases}$$

Se ve que las coordenadas experimentan una sustitución lineal. Designando por

$$\mathrm{X} = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} \quad \text{e} \quad \mathrm{Y} = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{bmatrix}$$

las matrices columna asociadas a las referencias de **v** y **w**, se obtiene inmediatamente la escritura matricial de la transformación

$$\mathrm{Y} = \mathrm{AX},$$

en la que A es la matriz anteriormente definida.

Rango.—Sabemos que el rango de la aplicación es la dimensión del subespacio de F engendrado por $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$; por consiguiente, es el rango de los vectores columna de A, es decir, el rango de esta matriz.

2. Ejemplos.—Como se sabe, la geometría ordinaria proporciona numerosos ejemplos de aplicaciones lineales (véase pág. 336). Tomemos el punto O, que queda invariante[1] en la transformación, como

[1] A este respecto, una traslación no es una transformación lineal, sino una transformación afín.

origen del sistema de coordenadas $Oxyz$. Si $M(x, y, z)$ es el punto inicial y $M'(x', y', z')$ su imagen, se debe verificar que

$$\begin{cases} x' = ax + by + cz, \\ y' = a'x + b'y + c'z, \\ z' = a''x + b''y + c''z, \end{cases} \qquad \text{o} \qquad X' = AX,$$

designando por A la matriz cuadrada de la transformación. La base común puede ser cualquiera; pero, en ciertas consideraciones métricas, la supondremos ortonormada.

a) Transformación idéntica:

$$x' = x, \qquad y' = y, \qquad z' = z.$$

La matriz es I.

b) Simetría con relación al origen: $x' = -x$, $y' = -y$, $z' = -z$; matriz $A = -I$.

c) Simetría oblicua con relación a Oz:

$$x' = -x, \qquad y' = -y, \qquad z' = z;$$

matriz diagonal

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

d) Homotecia de centro O *y razón* k:

$$x' = kx, \qquad y' = ky, \qquad z' = kz;$$

matriz escalar $A = kI$.

e) Proyección sobre el plano xOy: $x' = x$, $y' = y$, $z' = 0$;

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Los operadores anteriores se definen en geometría afín. La geometría métrica introduce otros nuevos; así, p. ej.,

f) Rotación de ángulo α *alrededor de* Oz:

$$\begin{cases} x' = x \cos \alpha - y \operatorname{sen} \alpha, \\ y' = x \operatorname{sen} \alpha + y \cos \alpha, \\ z' = z; \end{cases} \qquad A = \begin{bmatrix} \cos \alpha & -\operatorname{sen} \alpha & 0 \\ \operatorname{sen} \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

g) Semejanza directa que conserva Oz:

$$\begin{cases} x' = ax - by, \\ y' = bx + ay, \\ z' = z; \end{cases} \qquad A = \begin{bmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

EJERCICIOS:

1. En geometría afín y en un espacio de dos o tres dimensiones, descríbase la transformación geométrica asociada a cada una de las matrices

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ k & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix};$$

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. Determínese la matriz de cada una de las siguientes transformaciones geométricas, referidas a ejes rectangulares xOy.

a) Proyección del punto $M(x, y)$ sobre la bisectriz del primer cuadrante.

b) El punto M' es el simétrico de M con relación a la recta $y = 2x$.

c) El punto $M_1(1, -1)$ tiene por imagen $M_1'(0, 2)$ y al punto $M_2(2, 3)$ le corresponde el $M_2'(1, -1)$.

3. La transformación $x' = x + \lambda y$, $y' = y$ se llama deslizamiento paralelo a Ox. ¿Cómo se transforman los vectores situados sobre el eje Ox? Demuéstrese que los vectores situados en el eje Oy se transforman como si se tratase de una semejanza, de la que se determinará el ángulo de rotación y la razón de homotecia. La misma cuestión para los vectores situados en la bisectriz del primer cuadrante.

4. Consideramos la transformación $x' = x - y$, $y' = x + y$, referida a dos ejes rectangulares. Indíquese la construcción geométrica del punto $M'(x', y')$, a partir del punto $M(x, y)$.

5. Determínese la matriz de la transformación lineal que hace corresponder a los vectores $V(1, -1)$ y $V'(-3, 2)$ de R^2 los vectores $W(-1, -2, -5)$ y $W'(0, 5, 13)$ de R^3. ¿Cuál es el subespacio de R^3 engendrado por los transformados de los vectores de R^2?

6. Con relación a un triedro ortonormado $Oxyz$, determínese la matriz A del operador que asocia a todo punto M su simétrico respecto al plano P perpendicular al vector unitario $\mathbf{n}(\alpha, \beta, \gamma)$ trazado por O.

Determínese la matriz A' del operador que asocia a **OM** su proyección **OH** sobre la recta soporte del vector n trazada por O.

Compruébense geométricamente las siguientes relaciones:

$$A = I - 2A'; \qquad A'^2 = A'; \qquad A^2 = I.$$

7. Consideramos la transformación lineal $y_1 = x_2 + x_3$; $y_2 = x_1 + x_3$; $y_3 = x_1 + x_2$, referida a un triedro ortonormado.

Examínese cómo se transforman los vectores **OM** cuando M pertenece al plano $x_1+x_2+x_3=0$, o cuando está situado sobre la recta $x=y=z$. Dedúzcase de lo anterior una definición geométrica de la transformación.

8. En el espacio R^3 de la geometría clásica, hállese la matriz de un operador que deje invariante el plano xOy.

El mismo problema suponiendo, además, que el eje Oz es también un subespacio invariante.

9. En R^3 un operador lineal conserva Ox, así como el plano engendrado por los vectores (1, 2, 1), (0, 1, 2).

Escríbase la matriz eligiendo una base adaptada a los dos subespacios.

3. Isomorfismo entre matrices y operadores.—Por existir una correspondencia biunívoca entre el conjunto de las matrices (m, n) y el de las aplicaciones de E en F, sabemos que es posible definir operaciones en uno de ellos tomando las imágenes de las operaciones definidas sobre el otro, de manera que resultan isomorfos[1]. Dado el carácter de este libro nos limitaremos a comprobar que la biyección anterior establece un isomorfismo de las estructuras vectoriales de los dos conjuntos considerados. Es suficiente demostrar que se conservan las combinaciones lineales. Si T y T′ son dos aplicaciones, la aplicación $\lambda T+\mu T'$ hace corresponder a un vector cualquiera $\mathbf{e}_i \in E$, por ejemplo, el vector

$$(\lambda T+\mu T')\mathbf{e}_i=\lambda T\mathbf{e}_i+\mu T'\mathbf{e}_i$$

cuya componente de orden j es $\lambda a_i^j+\mu a_i'^j$, designando por a_i^j y $a_i'^j$ las componentes correspondientes de $T\mathbf{e}_i$ y $T'\mathbf{e}_i$. Volviendo a las matrices, se ha formado así la correspondiente combinación lineal de los términos homólogos de las matrices de T y T′. Por consiguiente, a $\lambda T+\mu T'$ corresponde la matriz $\lambda A+\mu A'$.

Ya que los dos espacios considerados son isomorfos, tienen la misma dimensión, mn, como se desprende de un resultado evidente para las matrices.

4. Composición de aplicaciones y producto de matrices.—Consideremos tres espacios vectoriales E_n, E_m, E_l; sea T una aplicación de E_n en E_m y T′ una aplicación de E_m en E_l. La compuesta $T'\circ T$ es una aplicación de E_n en E_l. Utilizando bases, se tienen las matrices

[1] Esto es lo que se hace muchas veces para definir las operaciones con las matrices a partir de las operaciones con las aplicaciones. Hemos preferido mantener dos estudios independientes, teniendo cada uno su propia justificación.

$A(m, n)$ y $B(l, m)$ tales que

$$\text{por T:} \quad Y = AX, \qquad \text{por T':} \quad Z = BY.$$

Eliminando Y, se tiene $Z = BAX$, de donde la matriz BA de la aplicación compuesta $T' \circ T$.

TEOREMA.—*En la biyección entre aplicaciones lineales y matrices de dimensiones convenientes, a la composición de dos aplicaciones corresponde la multiplicación de las matrices asociadas.*

OBSERVACIÓN.—Se puede emplear todavía aquí la composición de aplicaciones para definir el producto de matrices. De esta forma se evita la necesidad de demostrar ciertas propiedades de la multiplicación matricial.

Tenemos:

$$\text{por T:} \quad y^i = \sum_1^n a^i_\alpha x^\alpha; \qquad \text{por T':} \quad z^j = \sum_1^m b^j_\beta y^\beta.$$

Reemplazando las y, se deduce que

$$z^j = \sum_1^n \left(\sum_1^m b^j_\beta a^\beta_\alpha \right) x^\alpha.$$

La matriz asociada a $T' \circ T$ se obtiene, por consiguiente, haciendo el producto de filas por columnas de las matrices B y A; es la definición clásica.

5. Trasposición.—Sabemos que a toda aplicación T de E en F corresponde una aplicación 'T de F^* (dual de F) en E^* (dual de E). Por descontado, suponemos que E y E^*, así como F y F^*, están referidos a bases duales, $\{e_i\}$, $\{e^{*i}\}$ y $\{f_j\}$, $\{f^{*j}\}$. En estas condiciones, resulta la propiedad siguiente.

TEOREMA.—*Las matrices asociadas a dos aplicaciones traspuestas son traspuestas una de otra.*

Sea A la matriz de T; su columna *i-ésima* es la matriz del vector Te_i referida a la base $\{f_j\}$. La propiedad enunciada equivale a demostrar que las filas de A están formadas por las coordenadas de los diferentes vectores $'Tf^{*j} (j = 1, 2, \ldots, m)$ referidos a la base $\{e^{*i}\}$.

La *j-ésima* coordenada de $\mathbf{Te}_i$, o sea a_i^j, puede interpretarse como producto escalar (véase pág. 357):

$$a_i^j = \langle \mathbf{Te}_i, \mathbf{f}^{*j} \rangle.$$

Pero la definición de trasposición demuestra que

$$\langle \mathbf{Te}_i, \mathbf{f}^{*j} \rangle = \langle \mathbf{e}_i, {}^t\mathbf{Tf}^{*j} \rangle$$

y cabe interpretar el último producto como la coordenada de orden i del vector ${}^t\mathbf{Tf}^{*j}$ en la base $\{\mathbf{e}^{*i}\}$; esta coordenada es a_i^j. Dejando fijo j y haciendo variar i de 1 a n, se ve que las n coordenadas de ${}^t\mathbf{Tf}^{*j}$ constituyen la *j-ésima* fila de A (por consiguiente, la *j-ésima* columna de su traspuesta).

Consecuencia.—Sabemos que T y ${}^tT$ tienen el mismo rango. Por tanto, el rango de las columnas de una matriz (rango de T) es igual al de las filas (rango de ${}^tT$). Es el teorema del rango, que se vuelve a encontrar después de un largo rodeo; la demostración anterior apelaba solo a las nociones necesarias de dependencia e independencia lineales.

ENDOMORFISMOS DE UN ESPACIO VECTORIAL

6. Isomorfismo de anillos. Proseguimos el estudio de las aplicaciones lineales en el caso importante de las aplicaciones de un espacio en sí mismo, es decir, de los endomorfismos (pág. 343) de dicho espacio. Por razones de sencillez, suponemos que el original v y su imagen w = Tv están referidos a la misma base. Todo endomorfismo (operador) queda entonces caracterizado por una matriz cuadrada A, cuyas columnas son las diferentes coordenadas de los vectores $\mathbf{Te}_1$, $\mathbf{Te}_2$, ..., $\mathbf{Te}_n$, y la operación se escribe con notación matricial:

$$Y = AX \quad (X \text{ e } Y \text{ matrices columna de } v \text{ y } w).$$

Si los endomorfismos T y T' responden a las matrices A y B, respectivamente, su compuesto $T' \circ T$ responde a la matriz BA. El isomorfismo establecido (en cuanto a sus propiedades vectoriales) entre matrices cuadradas y operadores se conserva después de la introducción de las leyes multiplicativas; se obtiene finalmente un isomorfismo de anillos.

TEOREMA.—*El anillo de los endomorfismos de un espacio vectorial de n dimensiones es isomorfo al anillo de las matrices cuadradas de orden n. La correspondencia se establece después de elegir una base del espacio.*

7. Determinante de un endomorfismo.—Sea T un endomorfismo de un espacio E. Supongamos elegida una función determinante Δ no degenerada de E, y consideremos la aplicación de $(E)^n$ en K definida por

$$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) \longrightarrow \Delta(T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n).$$

De las propiedades de Δ y de T resulta inmediatamente que esta aplicación es n veces lineal y antisimétrica; por consiguiente, es una función determinante Δ_1, de donde

$$\Delta(T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n) = \Delta_1(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n). \qquad [1]$$

El teorema de la página 373 prueba que

$$\Delta_1(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = A\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n), \qquad [1\ bis]$$

donde A es una constante independiente de la elección de los vectores. Demostremos, además, que esta constante no depende de la elección de Δ. En efecto, el cambio de Δ en $\mu\Delta$ cambia Δ_1 en $\mu\Delta_1$ y la relación [1 *bis*] continúa siendo válida para las nuevas funciones.

Definición.—La constante A que, en definitiva, depende solo de la transformación T, se llama *determinante* de este endomorfismo.

Las relaciones [1] y [1 *bis*] se escriben entonces

$$\Delta(T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n) = (\det T)\Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n). \qquad [2]$$

TEOREMA I.—*El determinante de un endomorfismo es igual al determinante de la matriz de este endomorfismo en una base cualquiera.*

Sea la base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Sabemos que la matriz A de T admite por columnas sucesivas las coordenadas de $T\mathbf{e}_1, T\mathbf{e}_2, \ldots, T\mathbf{e}_n$ en la base considerada. Según la fórmula fundamental de la página 373,

$$\Delta(T\mathbf{e}_1, T\mathbf{e}_2, \ldots, T\mathbf{e}_n) = |A|\Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n).$$

Comparando con la fórmula [2], se ve que $\det T = |A|$.

TEOREMA II.—*El determinante del producto de dos endomorfismos*

T y T' *es igual al producto de los determinantes de estos endomorfismos.*

Se tiene:

$$\Delta(Tv_1, \ldots, Tv_n) = (\det T)\Delta(v_1, v_2, \ldots, v_n),$$

y

$$\begin{aligned}\Delta(T'Tv_1, \ldots, T'Tv_n) &= (\det T')\Delta(Tv_1, Tv_2, \ldots, Tv_n) \\ &= (\det T'T)\Delta(v_1, v_2, \ldots, v_n).\end{aligned}$$

Comparando estas igualdades, se ve que $(T'T) = \det T' \cdot \det T$.

OBSERVACIÓN.—Sabemos que T'T admite por matriz el producto BA de las matrices B y A de T' y T. Pasando a determinantes, se deduce que

$$|BA| = |B| \cdot |A|.$$

Este es el teorema de la página 378.

8. Grupo lineal.— Particularicemos los operadores considerando las biyecciones lineales de E_n en sí mismo; es decir, los automorfismos de E_n. Forman un *grupo lineal* (véase pág. 136), y a todo automorfismo T corresponde otro T^{-1} tal que

$$T^{-1} \circ T = T \circ T^{-1} = I.$$

En una base, las matrices A y B de estos operadores son tales que

$$BA = AB = I.$$

Son dos matrices inversas una de otra, y recíprocamente.

TEOREMA.—*El grupo de los automorfismos de un espacio vectorial (grupo lineal) es isomorfo al grupo de las matrices cuadradas regulares. La correspondencia se establece después de elegida una base.*

Una biyección de un espacio en sí mismo es una aplicación regular en el sentido dado en la página 339; el núcleo queda reducido al vector nulo. Para que suceda así es necesario y suficiente que su matriz en una base cualquiera sea regular; es decir, que su determinante sea distinto de cero.

Operador singular.—Por el contrario, el operador se llama singular cuando no se cumplen las condiciones anteriores. Existe, pues, un núcleo no reducido a 0 y la aplicación no es suprayectiva (véase pági-

na 344). Por tanto, un operador singular aplica de manera no biunívoca el espacio entero sobre uno de sus subespacios. Además, la suma de las dimensiones del núcleo y del subespacio imagen es igual a la dimensión del espacio inicial.

OBSERVACIONES.—1.ª Hemos encontrado ya el grupo lineal (página 136) al estudiar los cambios de base, pero bajo un aspecto diferente.

Cuando se trata de un operador (caso actual) un automorfismo T actúa sobre los vectores del espacio para transformarlos de acuerdo con

$$\mathbf{w} = T\mathbf{v} \qquad \text{o} \qquad Y = AX,$$

pero X e Y están referidos a la *misma base*, y los vectores de la base, como tales, no son transformados por el operador T.

Por el contrario, en un cambio de base, los nuevos vectores de la base son Te_i, mientras que los vectores del espacio permanecen fijos. La matriz X de un vector se convierte en X′ tal que $X = AX'$ (A matriz de cambio).

Decimos que se trata de dos aspectos complementarios de la teoría de las transformaciones regulares.

2.ª Además, las dos interpretaciones de la matriz inversa de A son diferentes. En el caso operador, A^{-1} es la matriz de T^{-1} en la base dada: las columnas son las coordenadas de los vectores $T^{-1}e_i$ en esta base.

Al contrario, en un cambio de base, A^{-1} es la matriz de la antigua base $\{e_i\}$ referida a la nueva $\{Te_i\}$. El lector precisará la razón por la que son iguales estas dos matrices.

9. Matrices semejantes.—La teoría anterior permite resolver fácilmente el siguiente problema.

Dado un operador lineal T, regular o no, representado en cierta base por la matriz A, se desea hallar la nueva matriz que lo caracteriza cuando se hace un cambio de base definido por la matriz de cambio P.

Sea $Y = AX$ la transformación considerada. Con relación a la nueva base, los vectores tienen por matrices $X_1 = P^{-1}X$ e $Y_1 = P^{-1}Y$. De ellas deducimos X e Y, que llevadas a la relación anterior la convierten en $PY_1 = A(PX_1)$; es decir, $Y_1 = (P^{-1}AP)X_1$.

Lo que demuestra que la nueva matriz es $B = P^{-1}AP$.

Definición.—Dos matrices cuadradas A y B que verifican la relación anterior se llaman *semejantes*. La relación «A es semejante a B» es, de acuerdo con su significado geométrico intrínseco, una relación de equivalencia definida sobre el conjunto de las matrices cuadradas de orden n (véase ejercicio 5, más adelante).

Una clase de semejanza está constituida por el conjunto de las matrices que corresponden a un mismo operador referido a todas las bases del espacio.

Recíprocamente, dadas dos matrices A y B, ¿bajo qué condiciones son semejantes? En la sección siguiente se dará una respuesta parcial a esta pregunta. Señalemos, no obstante, que una condición necesaria es la igualdad de los rangos de A y B, iguales al rango del operador.

Propiedades elementales de la semejanza.—1. Si A es semejante a B, y A_1 lo es a B_1, para la misma matriz de cambio, $\lambda A + \mu B$ es semejante a $\lambda A_1 + \mu B_1$, siendo λ y μ dos escalares del cuerpo considerado.

2. Si A es semejante a B, A^n y B^n son semejantes con la misma matriz de cambio, siendo n un entero relativo cualquiera. En efecto, de $B = P^{-1}AP$ se deduce que

$$B^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P,$$

la cual se generaliza para cualquier entero positivo, n. Finalmente, se tiene:

$$B^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}P.$$

3. Si $f(A)$ es un polinomio en A, de acuerdo con las propiedades anteriores tendremos:

$$f(B) = P^{-1}f(A)P.$$

En general, la aplicación de A en $P^{-1}AP$ es un automorfismo del anillo de las matrices cuadradas de orden n.

EJERCICIOS:

1. Estúdiese en R^3 el cambio de base definido por

$$E_1 = (1, 0, 1);\ E_2 = (0, 1, 1);\ E_3 = (1, 1, 0).$$

2. Indíquese una transformación lineal que aplique R^3 sobre el subespacio engendrado por los vectores (2, 1, 1) y (3, 0, −1).

3. En el espacio de la geometría corriente referido a un triedro ortonormado se considera la transformación T de matriz

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Demuéstrese que se trata de una transformación singular y precísese el subespacio engendrado por los transformados de los vectores del espacio. Caracterícese geométricamente esta transformación.

4. Se considera en un sistema de ejes rectangulares la transformación $x' = ax$, $y' = y$. Fórmese la matriz de este operador lineal cuando se toman como nuevos ejes las bisectrices de los antiguos.

5. Demuéstrese mediante el cálculo que la semejanza es una relación de equivalencia.

6. A una transformación regular T de matriz P se asocia un cambio de base. Demuéstrese que P es también la matriz de T en la nueva base.

10. Subespacio invariante.—Dado que para un endomorfismo se confunden el espacio inicial y el espacio imagen, se plantea la cuestión de la existencia de elementos invariantes (vectores, subespacios).

Definición.—Se dice que un subespacio E′ de E es *invariante* o *estable* para un operador T, si a todo $\mathbf{v} \in E'$, corresponde $T\mathbf{v} \in E'$. Así, por ejemplo, el núcleo es siempre un subespacio invariante, pero existen otros muchos.

Si consideramos E′ como espacio vectorial, se ve que T define en él un operador lineal, que se llama *inducido* por T en E′.

Precisemos la forma de la matriz adaptando a ella la base. Se eligen los p primeros vectores de la base en E′; si este subespacio es de p dimensiones lo completamos con otros $n-p$ vectores.

Como los transformados $T\mathbf{e}_1$, $T\mathbf{e}_2$, ..., $T\mathbf{e}_p$ pertenecen a E′, la matriz del operador en tal base tendrá sus p primeras columnas terminadas por $n-p$ ceros.

Se obtiene un caso más interesante cuando dos subespacios suplementarios E′ (p dimensiones) y E″ son invariantes por separado; T induce un operador T_1 en E′ y un operador T_2 en E″. El conjunto de estos operadores permite reconstituir T, puesto que, para todo $\mathbf{v} \in E$, se tiene:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2, \qquad \mathbf{v}_1 \in E', \qquad \mathbf{v}_2 \in E''$$

y

$$T\mathbf{v} = T_1\mathbf{v}_1 + T_2\mathbf{v}_2.$$

Adaptando la base a los dos subespacios, se ve que la matriz toma la forma siguiente $\begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix}$ (A_1 de orden p, A_2 de orden $n-p$), llamada, a veces, *suma directa* de A_1 y de A_2.

EJERCICIO:

Sea $P(x)$ un polinomio del anillo $K[x]$. Si T es un endomorfismo de un espacio E, se le asocia $T'=P(T)$. Demuéstrese que el núcleo de T' es un subespacio invariante para T.

DIAGONALIZACION DE UN OPERADOR

Sea T un operador lineal de matriz A en una base. Un cambio de base modifica la matriz, por lo que es natural hallar una base en la que la matriz de T sea lo más sencilla posible. Esa matriz pertenece a la clase de semejanza de A. Es la *reducida*, o matriz canónica, de esta clase. Para que dos matrices A y B sean semejantes es necesario y suficiente que las correspondientes reducidas sean idénticas.

La sencillez de los cálculos relativos a las matrices diagonales justifica la búsqueda de una reducida diagonal y vamos a indicar las condiciones necesarias y suficientes para que un operador sea *diagonalizable*. No estudiaremos la forma canónica que se debe adoptar para un operador no diagonalizable.

11. Vector propio y valor propio.—Sea T un operador representado por una matriz diagonal $A=\text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ en determinada base. La interpretación corriente de las columnas de A demuestra que

$$T(\mathbf{e}_1)=\lambda_1\mathbf{e}_1, \qquad T(\mathbf{e}_2)=\lambda_2\mathbf{e}_2, \ldots, T(\mathbf{e}_n)=\lambda_n\mathbf{e}_n.$$

Los transformados de los vectores de la base son, pues, homotéticos de estos vectores, lo que nos conduce a las definiciones siguientes.

Definiciones.—1.ª Se llama *vector propio* de un operador T todo vector $\mathbf{v}\neq\mathbf{0}$ tal que su transformado le sea proporcional: $T(\mathbf{v})=\lambda\mathbf{v}$, siendo λ un escalar bien determinado.

2.ª Se llama *valor propio* de un operador, todo escalar $\lambda\in K$ (cuerpo base) tal que exista al menos un vector propio que le corresponda. Pueden existir varios.

Estas definiciones no excluyen el caso $\lambda=0$, para el cual, $T(\mathbf{v})=\mathbf{0}$

y $\mathbf{v}$ pertenece al núcleo. Recíprocamente, todo vector del núcleo es propio y corresponde a $\lambda = 0$.

3.ª En virtud de las definiciones precedentes puede decirse que un operador es diagonalizable si existe una base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ formada por vectores propios. Siendo los valores propios correspondientes λ_1, λ_2, ..., λ_n, la matriz del operador es $A = \text{diag}\,\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ en esta base. Los valores $\lambda_1, \ldots, \lambda_n$ son elementos cualesquiera del cuerpo base, distintos o confundidos. Los siguientes teoremas precisan las propiedades de los valores propios y de los vectores propios asociados.

TEOREMA I.—*El conjunto de los vectores propios asociados a un mismo valor propio constituye un subespacio de* E.

Sea λ el valor propio. El conjunto de los vectores propios asociados contiene un número $1 \leqslant p \leqslant n$ de vectores independientes $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Se tiene así:

$$T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_p\mathbf{v}_p) = \sum_\alpha k_\alpha T\mathbf{v}_\alpha = \sum_\alpha k_\alpha \lambda v_\alpha = \lambda \sum_\alpha k_\alpha v_\alpha;$$

todo vector del subespacio engendrado por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ es, por consiguiente, propio y corresponde al valor λ.

TEOREMA II.—*Sean* $E_1, E_2, \ldots, E_p$ *los subespacios propios que corresponden a p valores propios distintos dos a dos* $\lambda_1, \lambda_2, \ldots, \lambda_p$. *La suma* $E_1 + E_2 + \ldots + E_p$ *de estos subespacios es directa.*

El teorema es cierto para $p = 2$, porque, si $\mathbf{v}$ es común a E_1 y a E_2, tenemos

$$T\mathbf{v} = \lambda_1\mathbf{v}, \qquad T\mathbf{v} = \lambda_2\mathbf{v}, \qquad \text{de donde} \qquad (\lambda_1 - \lambda_2)\mathbf{v} = \mathbf{0}.$$

Puesto que $\lambda_1 \neq \lambda_2$, deducimos que $\mathbf{v} = \mathbf{0}$. Supongamos, pues, que se verifique la propiedad para $p - 1$ subespacios propios y demostremos que es cierta para p. Sea $\mathbf{v}_i \in E_i$ (i de 1 a p). La suma es directa si, y solo si, la condición

$$\sum_1^p \mathbf{v}_i = \mathbf{0} \quad [1] \qquad \text{exige} \qquad \mathbf{v}_i = \mathbf{0}.$$

Ahora bien: la igualdad [1] implica, sucesivamente,

$$T\left(\sum_1^p \mathbf{v}_i\right) = \mathbf{0} \implies \sum_1^p T\mathbf{v}_i = \mathbf{0} \implies \sum_1^p \lambda_i\mathbf{v}_i = \mathbf{0}.$$

Además, [1] implica igualmente

$$\lambda_1\left(\sum_1^p \mathbf{v}_i\right)=\mathbf{0}$$

y, restando miembro a miembro del último término de la implicación anterior, resulta:

$$\sum_2^p(\lambda_i-\lambda_1)\mathbf{v}_i=\mathbf{0}.$$

En el primer miembro, los vectores $(\lambda_i-\lambda_1)\mathbf{v}_i$ son propios y pertenecen a E_2, E_3, ..., E_p. Según la hipótesis de recurrencia, se puede escribir

$$(\lambda_2-\lambda_1)\mathbf{v}_2=\mathbf{0},\quad (\lambda_3-\lambda_1)\mathbf{v}_3=\mathbf{0},\ \ldots,\ (\lambda_p-\lambda_1)\mathbf{v}_p=\mathbf{0}$$

y, puesto que todos los escalares son distintos de cero, se debe tener:

$$\mathbf{v}_2=\mathbf{0},\quad \mathbf{v}_3=\mathbf{0},\ \ldots,\ \mathbf{v}_p=\mathbf{0};$$

de donde resulta también $\mathbf{v}_1=\mathbf{0}$. El teorema es cierto para p subespacios.

TEOREMA FUNDAMENTAL.—*Para que un operador* T *de* E *sea diagonalizable, es necesario y suficiente que la suma directa de sus subespacios propios se confunda con* E.

La condición es necesaria, porque si T es diagonalizable, existe una base formada por vectores propios. Se consideran los valores propios distintos y se asocia a cada uno de ellos el subespacio propio engendrado por los vectores de la base correspondiente. La descomposición de E que así se obtiene es una suma directa [1] y puede denotarse con el signo $\oplus$ (véase pág. 122).

Recíprocamente, supongamos p subespacios propios E_i tales que

$$\mathbf{E}=\mathbf{E}_1\oplus\mathbf{E}_2\oplus\ldots\oplus\mathbf{E}_p.$$

Sean $\{\mathbf{e}_{i_1}\}$ una base de E_1 (i_1 de 1 a n_1, dimensión de E_1), $\{\mathbf{e}_{i_2}\}$ una base de E_2, ..., $\{\mathbf{e}_{i_p}\}$ una base de E_p. En virtud de la hipótesis, la reunión de estas bases constituye una base de E formada por vectores

[1] El lector demostrará fácilmente que no podrían existir otros vectores propios que los pertenecientes a los subespacios precedentes.

propios; T es, pues, diagonalizable. Además, en tal base la matriz es, evidentemente,

$$A = \text{diag}\,\{\lambda_1 \ldots, \lambda_2 \ldots, \lambda_p \ldots\},$$

donde λ_i, valor propio correspondiente a E_i, figura un número de veces igual a la dimensión de este subespacio.

12. Obtención de los valores propios y de los vectores propios. TEOREMA.—*Para que un escalar* λ *sea valor propio de* T *es necesario y suficiente que el operador* $T - \lambda I$ *sea singular.*

Si λ es valor propio, existirá $\mathbf{v} \neq \mathbf{0}$ tal que $T(\mathbf{v}) = \lambda\mathbf{v}$, o también

$$T(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0}, \quad \text{o sea} \quad (T - \lambda I)\mathbf{v} = \mathbf{0}.$$

El núcleo del operador $T - \lambda I$ comprende, pues, el vector $\mathbf{v} \neq \mathbf{0}$ al menos y $T - \lambda I$ no es regular. El recíproco es obvio.

Polinomio característico.—Para expresar la condición anterior adoptemos una base de E y sea A la matriz de T; la matriz de $T - \lambda I$ es $A - \lambda I$, la cual será singular si su determinante es nulo:

$$|A - \lambda I| = \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 - \lambda & \ldots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \ldots & a_n^n - \lambda \end{vmatrix} = 0.$$

Los valores propios son, por tanto, las raíces del *polinomio característico* obtenido al desarrollar este determinante. Este polinomio es, efectivamente, de grado n, siendo $(-1)^n$ el coeficiente del término de mayor grado. A cada una de sus raíces corresponde un vector propio al menos.

Puede ocurrir que el polinomio característico no tenga todas sus raíces en el cuerpo K, lo que nos llevará a tomar para cuerpo de los escalares un supercuerpo. Así, p. ej., si K es el cuerpo de los números reales, a veces nos vemos precisados a pasar al cuerpo de los complejos si deseamos obtener una descomposición completa del polinomio.

Invariancia.—Según su significado intrínseco el polinomio característico no debe modificarse cuando se reemplaza A por una cualquiera de las matrices semejantes, $B = P^{-1}AP$. En efecto:

$$B - \lambda I = P^{-1}AP - P^{-1}\lambda IP = P^{-1}(A - \lambda I)P$$

y, pasando a los determinantes,

$$|B-\lambda I| = |P|^{-1}\cdot|A-\lambda I|\cdot|P| = |A-\lambda I|.$$

Los coeficientes son, pues, los mismos: son invariantes en el grupo lineal. En particular, la función

$$a_1^1 + a_2^2 + \ldots + a_n^n$$

que, salvo el signo, es el coeficiente de λ^{n-1}, se llama *traza* de la matriz. El determinante $|A|$ es también un invariante.

13. Diagonalización: método práctico.—El teorema anterior conduce al método clásico de diagonalización. Se forma el polinomio característico y se hallan sus raíces en el cuerpo base considerado. A cada raíz λ_i corresponde un subespacio propio: el núcleo del operador $T-\lambda_i I$, que se puede determinar. La dimensión de tal subespacio satisface al siguiente teorema.

TEOREMA.—*A toda raíz λ_i de orden k del polinomio característico corresponde un subespacio propio cuya dimensión es, a lo sumo, igual a k.*

Supongamos, en efecto, que la dimensión fuera $k+1$. Se formaría una base de E eligiendo los $k+1$ primeros vectores de la base en el subespacio propio asociado a la raíz λ_i, y los otros, en un subespacio suplementario. En tal base, la matriz del operador admitiría la siguiente descomposición en cajas:

$$A = \left|\begin{array}{c|c} \lambda_i I_{k+1} & B_1 \\ \hline O & B_2 \end{array}\right|,$$

en la que I_{k+1} designa la matriz unidad de orden $k+1$. El polinomio característico correspondiente a tal matriz es divisible por $(\lambda-\lambda_i)^{k+1}$, por lo que λ_i sería raíz de orden superior a k.

Sin embargo, a toda raíz simple del polinomio corresponde un subespacio propio de dimensión igual a 1. En efecto, existe siempre un vector propio al menos.

Consecuencias. *a)* Es posible ahora precisar las condiciones prácticas de diagonalización. El polinomio característico debe descomponerse por completo en el cuerpo base y, si $\lambda_1, \lambda_2, \ldots, \lambda_p$ son sus distintas raíces, de órdenes respectivos $k_1, k_2, \ldots, k_p$, el subespacio propio correspondiente a la raíz λ_i debe ser de dimensión k_i.

En virtud del teorema anterior, esta es la única posibilidad para que la suma directa de los subespacios propios sea igual a E.

El siguiente ejemplo demuestra que no siempre es así; por consiguiente, no siempre es posible la diagonalización.

b) Sea T el operador de R^2 de matriz

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

en la base canónica. Ya que su polinomio característico es

$$|A - \lambda I| = (\lambda - 1)^2,$$

el único valor propio será $\lambda_1 = 1$. Sin embargo, el subespacio propio correspondiente tiene dimensión uno: es λe_1, y la diagonalización resulta imposible.

Caso particular. TEOREMA.—*Un operador cuyo polinomio característico admite n raíces distintas es diagonalizable.*

Hemos visto que a toda raíz simple corresponde un subespacio propio de dimensión igual a 1. Existen en este caso n raíces simples distintas 2 a 2 y, por consiguiente, n subespacios propios de una dimensión. Su suma es directa y, por tanto, se confunde con E.

EJEMPLO.—Pruébese que la transformación T de matriz

$$A = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

puede diagonalizarse. Determínense los vectores propios.

El polinomio característico es

$$\begin{vmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5).$$

Los valores propios son 2 y 5 y la matriz es semejante a

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}.$$

Se obtienen los vectores propios correspondientes a $\lambda = 2$ resolviendo el sistema

$$\begin{cases} 4x + 2y = 2x, \\ x + 3y = 2y, \end{cases}$$

de donde

$$x=1, \quad y=-1.$$

Para $\lambda=5$, se halla igualmente

$$x=2, \quad y=-1.$$

La matriz de cambio para obtener la diagonalización es, por tanto,

$$\mathrm{P}=\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \qquad \text{y} \qquad \mathrm{P}^{-1}=\frac{1}{3}\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}.$$

Dejamos al cuidado del lector la comprobación de que

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}=\mathrm{P}^{-1}\mathrm{AP}.$$

14. Ecuación verificada por un operador diagonalizable.—Demostremos que la condición de diagonalización puede escribirse en forma de una ecuación a la que satisface el operador (o su matriz), independientemente de la consideración de los subespacios propios.

TEOREMA.—*Para que un operador* T, *cuyos valores propios distintos son* $\lambda_1, \lambda_2, \ldots, \lambda_p$, *sea diagonalizable es necesario y suficiente que*

$$\varphi(\mathrm{T})=(\mathrm{T}-\lambda_1\mathrm{I})(\mathrm{T}-\lambda_2\mathrm{I})\ldots(\mathrm{T}-\lambda_p\mathrm{I}) \qquad [1]$$

sea el operador nulo.

Demostremos en primer lugar que, si T es diagonalizable, $\varphi(\mathrm{T})$ es tal que $\varphi(\mathrm{T})\mathbf{v}=\mathbf{0}$ para todo $\mathbf{v}$. Se puede escribir $\mathbf{v}=\sum_1^p \mathbf{v}_i$, con $\mathbf{v}_i \in \mathrm{E}_i$ (subespacio propio), y bastará probar que para cada $\mathbf{v}_i$ tenemos $\varphi(\mathrm{T})\mathbf{v}_i=\mathbf{0}$; ahora bien, como $\mathbf{v}_i$ pertenece a E_i, se tiene que $(\mathrm{T}-\lambda_i\mathrm{I})\mathbf{v}_i=\mathbf{0}$. Siendo permutables los diferentes factores de $\varphi(\mathrm{T})$, colocamos $(\mathrm{T}-\lambda_i\mathrm{I})$ en último lugar y deducimos que $\varphi(\mathrm{T})\mathbf{v}_i=\mathbf{0}$.

Recíprocamente, supongamos que existen p escalares λ_i, distintos dos a dos y tales que $\varphi(\mathrm{T})=\mathrm{O}$. Demostremos que T es diagonalizable, siendo los escalares anteriores los valores propios.

Consideremos los p polinomios

$$\begin{aligned} \varphi_1&=(x-\lambda_2)(x-\lambda_3)\ldots(x-\lambda_p), \\ \varphi_2&=(x-\lambda_1)(x-\lambda_3)\ldots(x-\lambda_p), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \varphi_p&=(x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_{p-1}). \end{aligned}$$

Son primos entre sí en su conjunto y, por el teorema de Bezout, existen p polinomios $X_1, X_2, \ldots, X_p$ tales que

$$\sum_{1}^{p} \varphi_i(x)X_i(x) = 1.$$

Pasando a los operadores, se deduce que

$$\sum_{1}^{p} \varphi_i(T)X_i(T) = I;$$

es decir, que para todo vector $\mathbf{v}$,

$$\sum_{1}^{p} \varphi_i(T)X_i(T)\mathbf{v} = \mathbf{v}. \qquad [2]$$

Consideremos los p operadores $\varphi_i(T)X_i(T)$. Si $\mathbf{v}$ describe E, $\varphi_i(T)X_i(T)\mathbf{v}$ describirá el subespacio imagen E_i del *i-ésimo* operador. Demostremos que para T, E_i es un subespacio propio que corresponde a λ_i. Tenemos, en efecto,

$$\varphi(T) = (T - \lambda_i I)\varphi_i(T) = O$$

y

$$(T - \lambda_i I)\varphi_i(T)X_i(T) = O;$$

es decir, que para $\mathbf{v}_i \in E_i$, $(T - \lambda_i I)\mathbf{v}_i = \mathbf{0}$.

En virtud del teorema fundamental los subespacios $E_1, E_2, \ldots, E_p$ tienen una suma directa, y según la fórmula [2] anterior esta suma es igual a E. Por consiguiente, el operador T es diagonalizable.

EJEMPLO.—Se considera el operador T, de R^4, de matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ a' & b & 2 & 0 \\ a'' & b' & c & 2 \end{bmatrix},$$

¿Qué condiciones deben satisfacer los coeficientes desconocidos para que este operador sea diagonalizable?

Los valores propios distintos son 1 y 2. Una condición necesaria y suficiente es, por tanto,

$$(A - I)(A - 2I) = O.$$

Se comprobará que es necesario y suficiente elegir $a = c = 0$.

EJERCICIOS:

1. Demuéstrese que las raíces características de una matriz simétrica de orden 2 son reales. ¿Cuál es la condición para que sean distintas?

2. Sea T un operador diagonalizable cuyos valores propios son distintos dos a dos. Demuéstrese que todo vector propio de T es también vector propio de cualquier operador T′ cuyo producto con T sea conmutativo.

3. Si A es una matriz columna de altura n, demuéstrese que

$$|\lambda I - A\,{}^tA| = \lambda^n - ({}^tAA)\lambda^{n-1},$$

donde ${}^tA$ es la traspuesta de A.

4. Fórmese el polinomio característico de la matriz

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Generalícese.

5. Estúdiense las raíces y los valores propios de las siguientes matrices y, si ello es posible, determínese una matriz P que las diagonalice:

$$\begin{bmatrix} 4 & 4 \\ 1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 2 & 5 \\ 4 & 3 \end{bmatrix}, \quad \begin{bmatrix} 5 & 3 \\ -8 & -6 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 2 & -3 \\ 2 & 2 & -6 \\ -2 & 2 & -6 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & -4 \\ 3 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ 2 & -2 & 4 \end{bmatrix}.$$

6. Las mismas cuestiones para las matrices

$$\begin{bmatrix} 7 & -12 & -2 \\ 3 & -4 & 0 \\ -2 & 0 & -2 \end{bmatrix}, \quad \begin{bmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{bmatrix}, \quad \begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{bmatrix}.$$

7. Pruébese que las matrices

$$\begin{bmatrix} 13 & 16 & 16 \\ -5 & -7 & -6 \\ -6 & -8 & -7 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix}, \quad \begin{bmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{bmatrix}$$

son, respectivamente, semejantes a

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

y hállese en cada caso una matriz de cambio.

8. Las mismas cuestiones para las matrices

$$\begin{bmatrix} 3 & 7 & -3 \\ -2 & -5 & 2 \\ -4 & -10 & 3 \end{bmatrix}, \qquad \begin{bmatrix} 8 & 30 & -14 \\ -6 & -19 & 9 \\ -6 & -23 & 11 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 1 & -1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{bmatrix},$$

semejantes, respectivamente, a

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

9. Sea la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Hállense sus valores propios. Demuéstrese que es diagonalizable. Determínese una matriz de cambio sencilla que permita la diagonalización.

Diagonalícense A^2 y A^{-1}.

10. Sea la matriz

$$A = \begin{bmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{bmatrix}.$$

Determínense sus valores propios. ¿Por qué no es diagonalizable? Demuéstrese que es semejante a

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

y dése una matriz de cambio, P.

11. Sean $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Hállense los valores propios y los vectores propios de A.

Compruébese que A es semejante a B. Hállese P.

12. *a)* Fórmense todas las matrices diagonales D de orden dado que verifiquen una ecuación matricial $P(D)=0$, donde $P(x)$ es un polinomio completamente descomponible en el cuerpo considerado.

b) Si T es una matriz regular cualquiera y D una de las matrices anteriores, pruébese que la matriz $A=T^{-1}DT$ verifica también la ecuación $P(A)=0$. En el caso en que $P(x)=0$ admita solo raíces simples, demuéstrese que se obtienen así todas las soluciones de $P(A)=0$.

c) En el cuerpo de los reales resuélvanse las ecuaciones

$$A^2-I=O, \qquad A^2-3A+2I=O, \qquad A^2-A=O.$$

Se hallarán las matrices de orden 2 ó 3 que verifiquen estas ecuaciones.

13. Hállense sobre el cuerpo de los reales las matrices A que satisfacen a $A^2+I=O$.

a) En R^2 demuéstrese que son semejantes a $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Asóciese a A una transformación T y pruébese que existe un vector V al menos tal que TV es independiente de V. Refiérase T a la base así formada.

b) En R^3 no existen soluciones reales. Complétese el par V, TV con otro vector de modo que los tres formen una base.

c) Estúdiese el caso de R^4.

14. Determínense en R^2 y en R^3 las matrices A tales que $A^2=O$ (sígase el mismo método que en el problema anterior).

15. Demuéstrese que, cualquiera que sea el número real a, las dos matrices siguientes son semejantes:

$$\begin{bmatrix} 4-a & 1 & -1 \\ -6 & -1-a & 2 \\ 2 & 1 & 1-a \end{bmatrix}, \qquad \begin{bmatrix} 1-a & 1 & 0 \\ 0 & 1-a & 0 \\ 0 & 0 & 2-a \end{bmatrix}.$$

EJERCICIOS SOBRE EL CAPITULO XVII

17-1. A todo vector de coordenadas x, y, z del espacio de tres dimensiones, R^3, se hace corresponder en el espacio de cuatro dimensiones, R^4, el vector cuyas coordenadas u, v, w, t vienen dadas por

$$\begin{cases} u=x+ay+2az, \\ v=ax+y+az, \\ w=2ax+2ay+z, \\ t=(2a+1)x+3ay+(2a+1)z, \end{cases}$$

donde a es un parámetro real dado.

Hállense los vectores de R^4 que son las imágenes de los vectores de R^3 por esta correspondencia.

1.° Demuéstrese que las coordenadas u, v, w, t de estos vectores satisfacen una relación independiente de a.

2.° Determínense los valores de a para los cuales los vectores que satisfacen la relación anterior no son todos imágenes de puntos de R^3.

3.º Para cada uno de los valores de a calculados en la segunda parte del problema dense las relaciones necesarias y suficientes a las que deben satisfacer u, v, w, t para que sean coordenadas de un vector de R^4, imagen de un vector de R^3.

17-2. En R^2 se considera un operador lineal que transforma

$$X=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{en} \quad Y=AX, \quad \text{con} \quad Y=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Hállese la matriz que transforma

$$X^{[2]}=\begin{bmatrix} x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \end{bmatrix} \text{en } Y^{[2]}=\begin{bmatrix} y_1^2 \\ \sqrt{2}y_1y_2 \\ y_2^2 \end{bmatrix}.$$

Desígnese por $A^{[2]}$ esta matriz. Establézcase para dos matrices cuadradas A y B la fórmula $(BA)^{[2]}=B^{[2]}A^{[2]}$. Se observará que $A^{[2]}$ está definida por la relación $Y^{[2]}=(AX)^{[2]}=A^{[2]}X^{[2]}$, válida para X arbitrario, y se introducirá a continuación la transformación $Z=BY$. Demuéstrese que $|A^{[2]}|=|A|^3$.

17-3. 1.º Determínese el rango de la transformación lineal T_m, donde m es un parámetro real,

$$T_m \begin{cases} X=mx+y+z, \\ Y=x+my+z, \\ Z=x+y+mz. \end{cases}$$

¿Existen vectores sobre los que $T_0(m=0)$ opera como una homotecia? Determínense la razón y los vectores correspondientes.

2.º Escríbase la transformación T_0^2, es decir, la transformación obtenida efectuando T_0 dos veces consecutivas; estúdiese la transformación T_0^k, obtenida efectuando k veces seguidas T_0; demuéstrese que, si I es la transformación idéntica, se puede escribir:

$$T_0^k=a_kT_0+b_kI,$$

siendo a_k y b_k enteros que satisfacen la relación de recurrencia, que se indicará; demuéstrese que los cocientes a_k/b_k y a_k/a_{k-1} tienen, para $k\to\infty$, límites, que deben calcularse; pruébese que se puede escribir:

$$a_k=\lambda 2^k+\mu(-1)^k, \qquad b_k=\lambda'2^k+(-1)^k\mu',$$

siendo λ, μ, λ', μ' constantes, que se indicarán; dedúzcase de ello una expresión de la suma de la serie [1]

$$I+tT_0+\frac{t^2}{2}T_0^2+\ldots+\frac{t^n}{n!}T_0^n+\ldots$$

---

[1] Véase ejercicio 14-2, pág. 331.

17-4. Una transformación lineal T depende de un parámetro t. En una base dada de R^2 se traduce por la matriz

$$A=\begin{bmatrix} t+1 & t^2-10 \\ 1 & t+3 \end{bmatrix}.$$

1.º Calcúlense en función de t los valores propios de A.

2.º En general, A es semejante a una matriz diagonal. ¿Existen casos de excepción? ¿Por qué razón?

3.º Demuéstrese que, en tal caso, A es semejante a una matriz B de la forma

$$B=\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

¿Cuál es entonces la naturaleza de T?

17-5. 1.º Calcúlense los valores propios y determínense los vectores propios de la matriz $A=\begin{bmatrix} -7 & -6 \\ 12 & 10 \end{bmatrix}$. Dedúzcase la matriz de cambio P que permite diagonalizar A y compruébese la relación $D=P^{-1}AP$. Se elegirá D de manera que el menor coeficiente esté en la primera fila, siendo P lo más sencilla posible con coeficientes enteros.

2.º Calcúlese A^2. ¿Cuáles son los valores propios y los vectores propios de A^2?

3.º En un plano referido a un sistema de ejes rectangulares se considera una sucesión de puntos M_j, definida por las relaciones de recurrencia

$$u_{j+1}=-7u_j-6v_j, \qquad v_{j+1}=12u_j+10v_j$$

entre las coordenadas (u_j, v_j) del punto M_j y las (u_{j+1}, v_{j+1}) de M_{j+1}, dándose el punto de partida M_1. Demuéstrese que, cualquiera que sea M_1, los puntos M_j están alineados y, salvo un caso de excepción, se alejan indefinidamente cuando j tiende hacia infinito. Calcúlese el límite de $-\dfrac{u_{j+1}^2+v_{j+1}^2}{u_j^2+v_j^2}$ cuando j tiende a infinito.

17-6. Sean T una aplicación lineal de un espacio E en sí mismo y v_1 un vector de E. Se considera la sucesión indefinida de vectores

$$v_1,\ v_2=T(v_1),\ v_3=T(v_2)=T^2(v_1),\ \ldots,\ v_k=T(v_{k-1})=T^{k-1}(v_1)\ldots$$

1.º Demuéstrese que el subespacio F engendrado por el sistema $\{v_1, v_2, \ldots, v_k, \ldots\}$ es invariante respecto a la operación T.

2.º Se supone que E es de dimensión n. Sea p el mayor entero tal que $\{v_1, v_2, \ldots, v_p\}$ sea libre. Demuéstrese que este sistema es una base de F. Si F coincide con E, se dice que la sucesión v_k engendra E. ¿Cuál es entonces el valor de p?

3.º Se supone que la matriz A de T en la base $\{e_1, e_2, \ldots, e_n\}$ de E es diagonal. Demuéstrese que, para que exista un vector v_1 que engendre E, es necesario y suficiente que los valores propios de T sean distintos dos a dos.

4.º Se supone que v_1 engendra E. ¿Cuál es la forma de la matriz A de T con relación a la base $\{v_1, v_2, \ldots, v_n\}$? Se designará por a_i al elemento situado en la fila i y en la última columna de A'. Demuéstrese que el polinomio característico de T es

$$P(\lambda)=(-1)^n(\lambda^n-a_n\lambda^{n-1}-a_{n-1}\lambda^{n-2}-\ldots-a_1).$$

Pruébese que la matriz obtenida reemplazando en $P(\lambda)$ la variable λ por la matriz A' de T es nula. Se demostrará que la aplicación lineal correspondiente a P(A') transforma cada uno de los vectores de la base $\{v_1, v_2, \ldots, v_n\}$ en el vector nulo.

CAPITULO XVIII

FORMAS BILINEALES Y FORMAS CUADRATICAS

APLICACIONES BILINEALES

1. Forma bilineal.—Sea E un espacio vectorial de n dimensiones sobre R (ó C). Precisemos, para el caso de una aplicación bilineal $f(\mathbf{v}, \mathbf{w})$, de $(E)^2$ en su cuerpo base, la expresión general obtenida en la página 366. Sean $\mathbf{v} = \sum x^\alpha \mathbf{e}_\alpha$ y $\mathbf{w} = \sum y^\beta \mathbf{e}_\beta$ las componentes de los vectores en una base cualquiera. Desarrollemos $f(\mathbf{v}, \mathbf{w})$ utilizando la bilinealidad:

$$f(\mathbf{v}, \mathbf{w}) = \sum_{\alpha,\beta} x^\alpha y^\beta f(\mathbf{e}_\alpha, \mathbf{e}_\beta) - \sum a_{\alpha\beta} x^\alpha y^\beta \qquad [1]$$

con $a_{\alpha\beta} = f(\mathbf{e}_\alpha, \mathbf{e}_\beta)$. Se obtiene una forma bilineal de las coordenadas (véase pág. 313), forma definida por los n^2 coeficientes $a_{\alpha\beta}$; en general, $a_{\alpha\beta} \neq a_{\beta\alpha}$ y la forma no es simétrica; hay que tener en cuenta el orden de los vectores considerados en un par $(\mathbf{v}, \mathbf{w})$.

Los coeficientes $a_{\alpha\beta}$ dependen de la elección de la base. Sin embargo, para definir intrínsecamente una función (aplicación) bilineal, basta elegir la forma que la expresa en una base determinada, es decir, los n^2 coeficientes $a_{\alpha\beta}$. Cuando se cambia de base, $f(\mathbf{v}, \mathbf{w})$ queda invariante, y debemos saber cómo se obtiene su nueva expresión en la nueva base, lo que haremos a continuación.

Notación matricial.—Muchas veces es cómodo emplear la notación matricial para tratar estas cuestiones. En una base dada, la forma puede caracterizarse por una matriz cuadrada $A = [a_{ij}]$ (i, índice de fila, j, índice de columna). Representado un vector por una matriz columna $X = [x^i]$, $Y = [y^j]$, en la expresión [1] reconocemos un producto de matrices:

$$f(\mathbf{v}, \mathbf{w}) = \sum_{\alpha\beta} x^\alpha (a_{\alpha\beta} y^\beta) = {}^tXAY,$$

siendo necesario usar la traspuesta de X en la escritura matricial. El

resultado es un escalar (una fila y una columna) y, por tanto, independiente de tal operación; es decir,

$$f(\mathbf{v}, \mathbf{w}) = {}^t\mathrm{Y}\,{}^t\mathrm{A}\mathrm{X}.$$

Esto equivale a caracterizar la forma por una nueva matriz ${}^t\mathrm{A} = [a_{ij}]$ (i, columna, j, fila).

Cambio de base.—Estudiemos el efecto de un cambio de base. Si P es la matriz de cambio, sabemos que las nuevas matrices de v y w son X' e Y' tales que $\mathrm{X} = \mathrm{PX}'$, $\mathrm{Y} = \mathrm{PY}'$. Se deduce así que

$$f(\mathbf{v}, \mathbf{w}) = ({}^t\mathrm{X}'\,{}^t\mathrm{P})\mathrm{A}(\mathrm{PY}') = {}^t\mathrm{X}'({}^t\mathrm{PAP})\mathrm{Y}';$$

la matriz asociada a la nueva forma es, por tanto, $\mathrm{A}' = {}^t\mathrm{PAP}$. Análogamente, a la matriz traspuesta ${}^t\mathrm{A}$ queda asociada la ${}^t\mathrm{A}'$. Se podrá, pues, escribir la expresión de $f(\mathbf{v}, \mathbf{w})$ en una base cualquiera.

EJERCICIOS:

1. Una aplicación bilineal de R^2 en R está definida por

$$f(\mathbf{e}_1, \mathbf{e}_1) = 1, \qquad f(\mathbf{e}_1, \mathbf{e}_2) = 1, \qquad f(\mathbf{e}_2, \mathbf{e}_1) = -1, \qquad f(\mathbf{e}_2, \mathbf{e}_2) = 3.$$

Determínese su matriz. Precísese el valor $f(\mathbf{v}, \mathbf{w})$ asociado a los vectores $\mathbf{v} = 2\mathbf{e}_1 - 3\mathbf{e}_2$, $\mathbf{w} = -\mathbf{e}_1 + 2\mathbf{e}_2$. Compárese con $f(\mathbf{w}, \mathbf{v})$.

Determínese la nueva matriz de esta aplicación cuando se toma por base $\mathbf{f}_1 = (1, 1)$, $\mathbf{f}_2 = (-1, 1)$.

2. Sea la aplicación f de R^2 en R definida en la base canónica por la forma

$$f(\mathbf{v}, \mathbf{w}) = 2x^1y^1 + x^1y^2 - x^2y^1 - 3x^2y^2 \quad [\mathbf{v} = (x^1, x^2),\ \mathbf{w} = (y^1, y^2)].$$

Hállese la nueva expresión de esta función cuando se elige la base $\mathbf{f}_1 = (2, 1)$, $\mathbf{f}_2 = (1, 1)$.

2. Aplicaciones asociadas a una función. Rango. TEOREMA.—*Toda función bilineal de un espacio vectorial define dos aplicaciones de este espacio* E *en un espacio dual* E*. *Estas aplicaciones son traspuestas una de la otra* [1].

Si suponemos en $f(\mathbf{v}, \mathbf{w})$, $\mathbf{v} = \mathbf{v}_1$, fijo, $f(\mathbf{v}_1, \mathbf{w})$ será una función lineal de w, que puede representarse como producto escalar de w por un

[1] Si φ es una aplicación de E en E* su traspuesta es también una aplicación de E en E* (véase ejercicio núm. 2, pág. 363).

vector determinado $\mathbf{v}_1^*$ del dual E^* de E:

$$f(\mathbf{v}_1, \mathbf{w}) = \langle \mathbf{w}, \mathbf{v}_1^* \rangle.$$

Como $\mathbf{v}_1$ puede elegirse arbitrariamente, se define así una aplicación φ de E en E^*: $\mathbf{v}_1 \longrightarrow \mathbf{v}_1^* = \varphi(\mathbf{v}_1)$. La linealidad de f con relación a $\mathbf{v}_1$ demuestra que φ es una aplicación lineal. El razonamiento se repite haciendo en $f(\mathbf{v}, \mathbf{w})$, $\mathbf{w} = \mathbf{w}_1$, fijo, resultando así una función lineal de v, es decir, un producto escalar por un vector dado $\mathbf{w}_1^*$ de E^*: $f(\mathbf{v}, \mathbf{w}_1) = \langle \mathbf{v}, \mathbf{w}_1^* \rangle$. Se obtiene así la aplicación lineal $\mathbf{w}_1 \longrightarrow \mathbf{w}_1^* = \varphi'(\mathbf{w}_1)$ de E en E^*.

Consideremos ahora $f(\mathbf{v}_1, \mathbf{w}_1)$. Según la primera fórmula se escribe:

$$f(\mathbf{v}_1, \mathbf{w}_1) = \langle \mathbf{w}_1, \varphi(\mathbf{v}_1) \rangle$$

y, por la segunda,

$$f(\mathbf{v}_1, \mathbf{w}_1) = \langle \mathbf{v}_1, \varphi'(\mathbf{w}_1) \rangle.$$

Para dos elementos $\mathbf{v}_1$ y $\mathbf{w}_1$ cualesquiera, se tiene así:

$$\langle \mathbf{v}_1, \varphi'(\mathbf{w}_1) \rangle = \langle \mathbf{w}_1, \varphi(\mathbf{v}_1) \rangle,$$

lo que demuestra que las aplicaciones φ y φ' son traspuestas una de la otra.

TEOREMA.—*Si* E *y* E^* *están referidos a bases duales, las matrices de las aplicaciones de* E *en* E^* *asociadas a la función* f *son idénticas a las matrices* A *y* ${}^t\mathrm{A}$ *de la forma bilineal en la base considerada.*

En una base, tenemos

$$f(\mathbf{v}, \mathbf{w}) = \sum_\alpha x^\alpha \left(\sum_\beta a_{\alpha\beta} y^\beta \right).$$

Supongamos $(y^1, y^2, \ldots, y^n)$ constantes y hagamos

$$y_\alpha^* = \sum a_{\alpha\beta} y^\beta;$$

por consiguiente,

$$f(\mathbf{v}, \mathbf{w}) = \sum_\alpha x^\alpha y_\alpha^*.$$

Se tiene un producto escalar referido a bases duales y, por consiguiente, las y_α^* son las coordenadas de $\mathbf{w}^*$ en la base de E^* dual de la

de E. Las fórmulas

$$y'_\alpha = \sum a_{\alpha\beta} y^\beta$$

son las ecuaciones de la transformación lineal designada anteriormente por φ'. Con las definiciones de la página 400, se ve que su matriz es A. Igualmente, la matriz de φ es ${}^tA$.

Rango.— Por definición, el rango común de estas aplicaciones es el rango de la función bilineal, el cual es también el rango de su matriz en una base cualquiera.

Si este rango es inferior a n, la función bilineal se llama *degenerada*. De manera más elemental, puede definirse el rango de una función bilineal como el rango de su matriz A en una base. En un cambio de base, pasamos de A a $A' = {}^tPAP$ y el rango se conserva porque P es regular (véase pág. 330). La definición tiene, pues, carácter intrínseco. Para que una aplicación sea degenerada es necesario y suficiente que sea nulo su discriminante, es decir, el determinante $|A|$ de su matriz en una base cualquiera.

EJERCICIOS:

1. Sea $f(\mathbf{v}, \mathbf{w}) = 2x^1y^1 - 4x^1y^2 + 5x^2y^1 + bx^2y^2$ una aplicación bilineal de $\mathbf{R}^2$ en R. Determínese b de manera que la forma sea degenerada. Precísese, entonces, el subespacio de los vectores $\mathbf{v}_0$ tales que $f(\mathbf{v}_0, \mathbf{w}) = 0$, cualquiera que sea $\mathbf{w}$, así como el subespacio de los vectores $\mathbf{w}_0$ tales que $f(\mathbf{v}, \mathbf{w}_0) = 0$, cualquiera que sea $\mathbf{v}$.

2. Generalícese la definición de los subespacios anteriores al caso de un espacio de n dimensiones. ¿Qué relación existe entre ellos?

APLICACIONES BILINEALES Y SIMETRICAS. FORMAS CUADRATICAS

3. Aplicaciones simétricas.—Aparecen muchas veces aplicaciones bilineales y simétricas: $f(\mathbf{v}, \mathbf{w}) = f(\mathbf{w}, \mathbf{v})$. En tal caso, la matriz A en una base cualquiera es simétrica, de donde

$$f(\mathbf{v}, \mathbf{w}) = {}^tXAY = {}^tYAX.$$

Lo que sigue de este capítulo se consagra únicamente al estudio de estas aplicaciones, lo que con frecuencia nos llevará a no mencionar la simetría. En este caso, las dos aplicaciones de E en E^* asociadas

con f coinciden: $\varphi = \varphi'$. Para $\mathbf{v}$ y $\mathbf{w}$ cualesquiera, se tiene:

$$f(\mathbf{v}, \mathbf{w}) = \langle \mathbf{v}, \varphi(\mathbf{w}) \rangle = \langle \mathbf{w}, \varphi(\mathbf{v}) \rangle.$$

El rango sigue siendo el de la aplicación φ; es decir, el de la matriz de f en una base cualquiera. Si es inferior a n, f se llama degenerada; ya que φ se efectúa entre dos espacios de la misma dimensión, resulta que su núcleo no se reduce solo al vector nulo. Existe, pues, un subespacio N de E tal que, si es $\mathbf{v}_1 \in \mathrm{N}$ y $\mathbf{w}$ arbitrario, se tiene:

$$f(\mathbf{v}_1, \mathbf{w}) = \langle \mathbf{w}, \varphi(\mathbf{v}_1) \rangle = 0.$$

Designaremos también por N el *núcleo* de la función bilineal. Un vector cualquiera del núcleo y un vector arbitrario de E dan un valor nulo a la función. En particular, para un vector del núcleo $f(\mathbf{v}_1, \mathbf{v}_1) = 0$.

De manera más elemental, obsérvese que la condición ${}^t\mathrm{YAX} = 0$, que suponemos verificada para Y arbitrario, exige que $\mathrm{AX} = 0$. El conjunto de los X que satisfacen esta condición define el núcleo, el cual, si A es de rango r, será un subespacio de $n - r$ dimensiones de E.

Cambio de base.—Si es P la matriz de cambio, la nueva matriz de la función será $\mathrm{A}' = {}^t\mathrm{PAP}$ y ${}^t\mathrm{A}' = \mathrm{A}'$. La relación anterior entre dos matrices A y A′ es una relación de equivalencia[1] (compruébelo el lector), justificada por el hecho de que ambas corresponden a la misma función. En una clase de equivalencia cabe hallar una reducida de forma sencilla; demostraremos que existe una infinidad de reducidas diagonales.

4. Forma cuadrática asociada.—Si suponemos $\mathbf{w} = \mathbf{v}$ en la función f, se obtiene una función de $\mathbf{v}$ denotada por $f_c(\mathbf{v})$: es la forma cuadrática asociada a f. Referida a una base es un polinomio homogéneo de segundo grado de las coordenadas del vector $\mathbf{v}$:

$$f_c(\mathbf{v}) = \sum_{\alpha\beta} a_{\alpha\beta} x^\alpha x^\beta = {}^t\mathrm{XAX}.$$

Tal polinomio define una *forma cuadrática* del vector $\mathbf{v}$. Para no hacer intervenir la referencia a una base, damos a continuación una definición intrínseca de la misma.

[1] En este caso, se dice *congruencia*.

Definición.— Se llama forma cuadrática[1] en un espacio E (R ó C) una aplicación f_c de E en R (ó C) que verifica la propiedad característica

$$f_c(\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda^2 f_c(\mathbf{v}) + 2\lambda\mu f(\mathbf{v}, \mathbf{w}) + \mu^2 f_c(\mathbf{w}), \qquad [1]$$

donde $f(\mathbf{v}, \mathbf{w})$ designa cierta función bilineal, cuyas relaciones con f_c vamos a precisar.

Haciendo $\lambda = \mu = 1$ en [1] y comparando $f_c(\mathbf{v} + \mathbf{w})$ con $f_c(\mathbf{w} + \mathbf{v})$, se ve que f es necesariamente simétrica. Si se hace $\mu = 0$, se obtiene $f_c(\lambda \mathbf{v}) = \lambda^2 f_c(\mathbf{v})$; finalmente, haciendo $\lambda = \mu = 1$ y $\mathbf{v} = \mathbf{w}$, resulta:

$$f_c(2\mathbf{v}) = 2f_c(\mathbf{v}) + 2f(\mathbf{v}, \mathbf{v});$$

pero $f_c(2\mathbf{v}) = 4f_c(\mathbf{v})$ y, por consiguiente, $f_c(\mathbf{v}) = f(\mathbf{v}, \mathbf{v})$; la forma cuadrática se deduce por el proceso anterior de la forma bilineal $f(\mathbf{v}, \mathbf{w})$, la cual recibe el nombre de *forma polar* de la forma cuadrática.

OBSERVACIÓN.—La fórmula $f_c(\mathbf{v}) = f(\mathbf{v}, \mathbf{v})$ permite obtener los valores de la forma cuadrática a partir de los de la forma bilineal. Pero, recíprocamente, si hacemos en [1] $\lambda = \mu = 1$, se tiene:

$$f(\mathbf{v}, \mathbf{w}) = \frac{1}{2}[f_c(\mathbf{v} + \mathbf{w}) - f_c(\mathbf{v}) - f_c(\mathbf{w})],$$

lo que prueba que se pueden calcular los valores de la forma bilineal a partir de los de la forma cuadrática asociada.

En particular, si suponemos $f_c(\mathbf{v}) = 0$ para todo vector de un subespacio, tendremos que $f(\mathbf{v}, \mathbf{w}) = 0$ para todo par de vectores de este subespacio.

EJERCICIOS:

1. Sea, en $\mathbf{R}^2$, la forma cuadrática

$$f_c(\mathbf{v}) = 2(x^1)^2 - 3x^1x^2 + 4(x^2)^2.$$

Determínese su nueva expresión cuando se toma por base $\mathbf{f}_1 = (1, 1)$, $\mathbf{f}_2 = (-1, 1)$. Escríbase la forma bilineal asociada.

2. Determínese en la base natural de $\mathbf{R}^3$ la matriz de una forma bilineal simétrica f tal que, para $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (-1, 2, 0)$, $\mathbf{v}_3 = (1, 0, 1)$, se tenga:

$$f(\mathbf{v}_1, \mathbf{v}_2) = 0, \quad f(\mathbf{v}_2, \mathbf{v}_3) = 4, \quad f(\mathbf{v}_1, \mathbf{v}_3) = -1,$$
$$f_c(\mathbf{v}_1) = 5, \quad f_c(\mathbf{v}_2) = 1, \quad f_c(\mathbf{v}_3) = 0.$$

[1] Sería más correcto decir función cuadrática o aplicación cuadrática, siendo la forma la expresión de esta función en una base determinada. Pero nos atendremos aquí al lenguaje usual.

3. Sea f_c la forma cuadrática asociada a una forma bilineal, f. Demuéstrese que

$$f_c(\mathbf{u}+\mathbf{v}+\mathbf{w})=f_c(\mathbf{u}+\mathbf{v})+f_c(\mathbf{v}+\mathbf{w})+f_c(\mathbf{w}+\mathbf{u})-f_c(\mathbf{u})-f_c(\mathbf{v})-f_c(\mathbf{w}).$$

5. Vectores conjugados. **Definición.**—1. Sea f una función bilineal simétrica. Se llaman conjugados en la aplicación dos vectores $\mathbf{v}$ y $\mathbf{w}$ tales que $f(\mathbf{v}, \mathbf{w})=0$. Esta relación es recíproca.

2. El núcleo de la función bilineal es el conjunto de los vectores conjugados de todos los vectores del espacio; para una función degenerada, solo el vector nulo verifica esta propiedad.

3. Un vector es conjugado de un subespacio si es conjugado de todos los vectores de este subespacio. Para ello es necesario y suficiente que sea conjugado de los vectores de una base del subespacio.

4. Finalmente, dos subespacios son conjugados si todo vector de uno de ellos es conjugado de todo vector del otro. En particular, si $\mathbf{v}$ y $\mathbf{w}$ son conjugados, también lo son los subespacios $\lambda\mathbf{v}$ y $\mu\mathbf{w}$.

TEOREMA 1.—*Si la forma cuadrática asociada a f toma un valor no nulo para un vector $\mathbf{v}$, todo vector $\mathbf{w}$ es la suma de un vector proporcional a $\mathbf{v}$ y de otro vector conjugado de $\mathbf{v}$. Esta descomposición es única.*

Si $\mathbf{w}=\lambda\mathbf{v}+\mathbf{w}_1$ es una descomposición del espacio buscado, se debe tener $f(\mathbf{v}, \mathbf{w}_1)=0$, lo que se escribe:

$$f(\mathbf{v}, \mathbf{w}-\lambda\mathbf{v})=f(\mathbf{v}, \mathbf{w})-\lambda f_c(\mathbf{v})=0,$$

de donde

$$\lambda=\frac{f(\mathbf{v}, \mathbf{w})}{f_c(\mathbf{v})}$$

Este valor queda determinado de manera única, de donde se deduce que $\mathbf{w}_1=\mathbf{w}-\lambda\mathbf{v}$.

TEOREMA 2.—*Todo subespacio de p ($p\leqslant n$) dimensiones de E puede referirse a una base formada por vectores conjugados dos a dos.*

Se busca en el subespacio E_p un vector $\mathbf{e}_1$ tal que $f_c(\mathbf{e}_1)\neq 0$. Si no existiera, se tendría $f_c(\mathbf{v})=0$ para todo $\mathbf{v}\in E_p$ y, por consiguiente (observación hecha en el apartado 4), todos los vectores serían conjugados dos a dos, y cualquier base de E_p verificaría la condición pedida.

En caso contrario, se elige una base $(\mathbf{e}_1, \ldots, \mathbf{e}_p)$ tal que $f_c(\mathbf{e}_1)\neq 0$. Se reemplaza todo $\mathbf{e}_i$ (i de 2 a p) por un vector $\mathbf{e}'_i=\mathbf{e}_i-\lambda\mathbf{e}_1$ conjugado

de $\mathbf{e}_1$. Los vectores $\{\mathbf{e}_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_p\}$ forman una base tal que $\mathbf{e}_1$ es conjugado del subespacio E'_{p-1} engendrado por $\{\mathbf{e}'_2, \ldots, \mathbf{e}'_p\}$.

Así se continúa aplicando el mismo método a E'_{p-1}: se busca $\mathbf{f}_2$ tal que $f_c(\mathbf{f}_2) \neq 0$. En el caso de no existir, toda base de E'_{p-1} completada con $\mathbf{e}_1$ responde al teorema; si no, tenemos $\{\mathbf{f}_2, \ldots, \mathbf{f}_p\}$ que se reemplaza por $\{\mathbf{f}_2, \mathbf{f}'_3, \ldots, \mathbf{f}'_p\}$ tal que el subespacio engendrado por $\{\mathbf{f}'_3, \ldots, \mathbf{f}'_p\}$ sea conjugado de $\mathbf{f}_2$.

Después de un número de operaciones análogas, inferior o igual a $p-1$, se obtiene con certeza una base que responde a las condiciones impuestas; sea $\{\mathbf{e}_1, \mathbf{f}_2, \mathbf{g}_3, \ldots, \mathbf{l}_p\}$. En general, hay una indeterminación en la elección de $\mathbf{e}_1$, $\mathbf{f}_2$, ... y existe una infinidad de bases que responden a la cuestión.

6. Diagonalización de una forma.—El teorema anterior puede aplicarse al espacio E entero. Resulta de ello que siempre es posible encontrar al menos una base $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ formada por vectores conjugados dos a dos: $f(\mathbf{e}_i, \mathbf{e}_j) = 0$. La matriz correspondiente, D, de la función es diagonal, siendo los elementos diagonales $f_c(\mathbf{e}_i)$:

$$D = \text{diag}\,\{f_c(\mathbf{e}_1), f_c(\mathbf{e}_2), \ldots, f_c(\mathbf{e}_n)\}.$$

El número de términos no nulos de la diagonal es igual al rango de D, es decir, al rango de la función; por consiguiente, es un número r bien determinado. Cualquiera que sea la matriz diagonal obtenida, tendrá el mismo número de términos no nulos en la diagonal.

7. Descomposición de una forma cuadrática en cuadrados.—El problema de la diagonalización se ha tratado desde un punto de vista intrínseco, pero lo podemos traducir con ayuda del cálculo sobre las formas. Sea

$$f_c(\mathbf{v}) = \sum a_{\alpha\beta} x^\alpha x^\beta$$

una forma cuadrática en una base determinada. Si es P la matriz de cambio de esta base a otra que asegura la diagonalización, se tiene para un vector cualquiera $X = PX'$, de donde $X' = P^{-1}X$. La forma $f_c(\mathbf{v})$ se convierte, en esta base, en

$$f_c(\mathbf{v}) = d_1(x'^1)^2 + d_2(x'^2)^2 + \ldots + d_r(x'^r)^2,$$

llamando r al rango. Las coordenadas x'^1, x'^2, ..., x'^r son formas lineales independientes en x^1, x^2, ..., x^n. La diagonalización consiste, pues,

en expresar $f_c(\mathbf{v})$ como suma de r cuadrados de formas lineales independientes en $x^1, x^2, \ldots, x^n$.

TEOREMA. — *Existe al menos una sustitución lineal regular* $x^i - \sum_p \alpha_p^i x'^p$ *que reduce una forma cuadrática cualquiera a una suma de r cuadrados, siendo r el rango de la forma.*

Método de Gauss.—El método del apartado 6 resuelve por completo el problema de la diagonalización. Sin embargo, el método de Gauss permite, calculando directamente sobre la forma cuadrática, obtener una solución particular, que determina el rango de la forma y su signatura (véase más adelante). Es de señalar que una forma cuadrática es un polinomio de segundo grado (eventualmente de primero) con relación a cada variable; así, p. ej., para x^1:

$$f_c(x^1, x^2, \ldots, x^n) = a(x^1)^2 + 2\varphi_1 \cdot x^1 + g_c(x^2, \ldots, x^n)$$
$$= a\left(x^1 + \frac{\varphi_1}{a}\right)^2 + g_c - \frac{\varphi_1^2}{a}$$

Puesto que $g_c - \dfrac{\varphi_1^2}{a}$ es una forma cuadrática de $n-1$ variables [1], el teorema queda demostrado, ya que es cierto para $n-1$: sabemos expresar f_c como suma de cuadrados de formas lineales. La independencia de estas se obtiene también por recurrencia, porque el primer cuadrado es el único que introduce la variable x^1.

Si f_c no contiene términos cuadrados, se ordena con relación a dos variables, x^1 y x^2:

$$f_c(x^1, x^2, \ldots, x^n) = ax^1x^2 + x^1\varphi_1 + x^2\varphi_2 + g_c(x^3, \ldots, x^n)$$
$$= a\left(x^1 + \frac{\varphi_2}{a}\right)\left(x^2 + \frac{\varphi_1}{a}\right) + g_c - \frac{\varphi_1\varphi_2}{a}$$
$$= \frac{a}{4}[(x'^1)^2 - (x'^2)^2] + g_c - \frac{\varphi_1\varphi_2}{a}$$

Se introducen así dos cuadrados, pero la forma restante es de $n-2$ variables.

[1] Suponemos $a \neq 0$; φ_1 es una forma lineal en $x^2, \ldots, x^n$, y g_c una forma cuadrática de las mismas variables.

EJEMPLO.—Tenemos, p. ej.,

$$f_c(\mathbf{v}) = x^2 - y^2 + 3z^2 + 4xy + 6xz + 8xz =$$

$$= (x + 2y + 3z)^2 \quad 5\left(y + \frac{2}{5}z\right)^2 - \frac{26}{5}z^2.$$

EJERCICIOS:

1. Descompónganse en cuadrados las siguientes formas cuadráticas:

a) $(x^1)^2 + 2x^1x^2 + 2(x^2)^2 + 4x^2x^3 + 5(x^3)^2$;
b) $5(x^1)^2 - 3(x^2)^2 + (x^3)^2 + 12x^1x^2 - 6x^1x^3 - 8x^2x^3$;
c) $x^1x^2 + x^2x^3 + x^3x^1$;
d) $(x^1)^2 + (x^2)^2 + (x^3)^2 - 2(x^4)^2 - 2x^1x^2 + 2x^1x^3 - 2x^1x^4 + 2x^2x^3 - 4x^2x^4$.

2. Sea la forma cuadrática que en una base de R^n se escribe

$$f_c(\mathbf{v}) = \sum_1^n (x^\alpha)^2 + \sum_{\alpha<\beta} x^\alpha x^\beta.$$

Demuéstrese que una descomposición en cuadrados de f es

$$f_c(\mathbf{v}) = (y^1)^2 + \frac{3}{4}(y^2)^2 + \frac{4}{6}(y^3)^2 + \ldots + \frac{n+1}{2n}(y^n)^2$$

y precísese la sustitución sobre las variables x^i.

FORMAS CUADRATICAS SOBRE EL CUERPO DE LOS NUMEROS REALES

8. Caso del cuerpo R. Ley de Sylvester.—En R es posible precisar la forma canónica de la matriz diagonal. Los coeficientes diagonales $d_i = f_c(\mathbf{e}_i)$ pueden ser positivos, negativos o nulos. Supondremos que, mediante una sustitución sobre los vectores que constituyen la base, los coeficientes positivos aparecen los primeros (numerados de 1 a p), a continuación los negativos (numerados de $p+1$ a r) y, finalmente, los coeficientes nulos ($r+1$ a n). Se toma una nueva base tal que

$$\mathbf{f}_i = \frac{1}{\sqrt{\pm d_i}}\mathbf{e}_i, \qquad (+ \text{ de 1 a } p, \quad - \text{ de } p+1 \text{ a } r); \qquad \mathbf{f}_i = \mathbf{e}_i \ (i > r).$$

Tenemos así:

$$f_c(\mathbf{f}_i) = 1 \ (1 \text{ a } p), \quad f_c(\mathbf{f}_i) = -1 \ (p+1 \text{ a } r), \quad f_c(\mathbf{f}_i) = 0 \ (r+1 \text{ a } n);$$

la matriz es

$$D = \text{diag } \{1, 1, \ldots, 1, -1, -1, \ldots, -1, 0, \ldots, 0\}$$

y la forma cuadrática correspondiente

$$f_e(\mathbf{v}) = (z^1)^2 + (z^2)^2 + \ldots + (z^p)^2 - (z^{p+1})^2 - \ldots - (z^r)^2, \qquad [1]$$

con $\mathbf{v} = \sum_1^n z^i \mathbf{f}_i$.

Ley de inercia de Sylvester.—*Todas las reducidas diagonales de una función bilineal y simétrica contienen el mismo número de términos positivos.*

En la base anterior tenemos la expresión [1] de f_e con p términos positivos. Supongamos que en otra base $\{\mathbf{g}_i\}$, la reducida tenga $p' < p$ términos positivos. Para $\mathbf{v} = \sum_1^n t^i \mathbf{g}_i$ se tendría

$$f_e(\mathbf{v}) = (t^1)^2 + \ldots + (t^{p'})^2 - (t^{p'+1})^2 - \ldots - (t^r)^2.$$

La primera expresión prueba que $f_e(\mathbf{v}) \geqslant 0$ para todos los vectores del subespacio de $n-(r-p)$ dimensiones engendrado por $\{\mathbf{f}_1, \ldots, \mathbf{f}_p, \mathbf{f}_{r+1}, \ldots, \mathbf{f}_n\}$. La segunda demuestra que $f_e(\mathbf{v}) < 0$ (en sentido estricto) para todos los vectores del subespacio de $r-p'$ dimensiones engendrado por $\mathbf{g}'_{p'+1}, \mathbf{g}'_{p'+2}, \ldots, \mathbf{g}'_r$. La suma de las dimensiones

$$n-(r-p)+r-p' = n+p-p' > n$$

y los dos subespacios no pueden ser disjuntos. Ahora bien: existe contradicción para los vectores de la interseccion; la hipótesis $p' > p$ llevaría a una contradicción análoga, por lo que $p = p'$.

Signatura.—Toda reducida diagonal de la clase de equivalencia de matrices asociada a una función bilineal dada comprende p números positivos y $r-q$ negativos, si r es el rango. El símbolo

$$[+ + \ldots +, - \ldots -] \qquad (p \text{ signos } +, r-q \text{ signos } -)$$

se llama *signatura* de la función bilineal.

TEOREMA.—*Dos funciones bilineales tienen la misma signatura si, y solo si, existe un automorfismo del espacio* E *que asocia los vectores*

de manera que las formas toman los mismos valores para vectores asociados.

Es suficiente razonar con los valores de las formas cuadráticas. Si f_c y g_c tienen la misma signatura, existe una base $\{\mathbf{f}_i\}$ y otra $\{\mathbf{g}_i\}$ de E tales que, si

$$\mathbf{v} = \sum x^i \mathbf{f}_i \qquad \text{y} \qquad \mathbf{w} = \sum x^i \mathbf{g}_i,$$

se tenga:

$$f_c(\mathbf{v}) = (x^1)^2 + \ldots + (x^p)^2 - (x^{p+1})^2 - \ldots - (x^r)^2;$$

$g_c(\mathbf{w})$ tiene la misma expresión. Por consiguiente, el automorfismo asociado a las dos bases es tal que $f(\mathbf{v}, \mathbf{v}') = g(\mathbf{w}, \mathbf{w}')$ si a $\mathbf{v}$ corresponde $\mathbf{w}$ y a $\mathbf{v}'$, $\mathbf{w}'$.

El recíproco se demuestra fácilmente considerando los vectores de la base.

9. Carácter de definición. Desigualdad de Schwartz.– Sea $f_c(\mathbf{v})$ la forma cuadrática en R asociada a una función bilineal. En ciertos casos, los valores que toma esta forma poseen propiedades características que vamos a precisar.

Forma definida positiva.— Supongamos que para todo $\mathbf{v}$ sea $f_c(\mathbf{v}) \geqslant 0$, y que la igualdad $f_c(\mathbf{v}) = 0$ exija $\mathbf{v} = \mathbf{0}$. El núcleo se reduce, ciertamente, al vector nulo, y la función es regular. La forma se llama *definida positiva*; para que así suceda es necesario y suficiente que su signatura comprenda n signos $+$.

Demostremos que en estas condiciones se verifica la *desigualdad de Schwartz:*

$$[f(\mathbf{v}, \mathbf{w})]^2 \leqslant f_c(\mathbf{v}) f_c(\mathbf{w}).$$

En efecto: hemos visto que

$$f_c(\lambda \mathbf{v} + \mathbf{w}) = \lambda^2 f_c(\mathbf{v}) + 2\lambda f(\mathbf{v}, \mathbf{w}) + f_c(\mathbf{w}).$$

Al variar λ, permaneciendo constantes $\mathbf{v}$ y $\mathbf{w}$, se verifica siempre que $f_c \geqslant 0$; por consiguiente, el trinomio del segundo miembro en λ tiene discriminante negativo o nulo, de donde

$$[f(\mathbf{v}, \mathbf{w})]^2 \leqslant f_c(\mathbf{v}) f_c(\mathbf{w}).$$

En esta relación, la igualdad exige que $\mathbf{v}$ y $\mathbf{w}$ sean dependientes.

En efecto: cabe suponer que uno de los vectores, $\mathbf{v}$, es no nulo (en caso contrario tendríamos la igualdad y el sistema sería dependiente) y hemos visto que (véase pág. 431)

$$\mathbf{w} = \mathbf{w}_1 + k\mathbf{v}, \qquad \text{con} \qquad f(\mathbf{v}, \mathbf{w}_1) = 0,$$

de donde

$$k^2[f_c(\mathbf{v})]^2 = f_c(\mathbf{v})f_c(\mathbf{w}_1) + k^2[f_c(\mathbf{v})]^2.$$

Reduciendo queda $f_c(\mathbf{w}_1) = 0$ [porque $f_c(\mathbf{v}) \neq 0$] y, por consiguiente, $\mathbf{w}_1 = \mathbf{0}$: los dos vectores $\mathbf{v}$ y $\mathbf{w}$ son proporcionales. El estudio de estas formas bilineales es importante y lo continuaremos en el próximo capítulo.

Forma semidefinida positiva.— Para todo $\mathbf{v} \in E$ se tiene $f(\mathbf{v}) \geqslant 0$, con $f_c(\mathbf{v}) = 0$ para ciertos vectores no nulos. La forma se llama *semidefinida positiva*. La desigualdad de Schwartz sigue siendo válida, como se comprueba inmediatamente. Demostremos que el rango de tal aplicación es inferior a n. En efecto, si $\mathbf{v}_1$ es tal que $f_c(\mathbf{v}_1) = 0$, y $\mathbf{w}$ es cualquiera, tenemos, en virtud de la desigualdad de Schwartz,

$$[f(\mathbf{v}_1, \mathbf{w})]^2 \leqslant f_c(\mathbf{v}_1)f_c(\mathbf{w}) = 0$$

y, por consiguiente, $f(\mathbf{v}_1, \mathbf{w}) = 0$ para $\mathbf{w}$ arbitrario; por tanto, $\mathbf{v}_1 \neq \mathbf{0}$ pertenece al núcleo.

La signatura de tal función no puede contener el signo $-$, ya que en ese caso existiría al menos un vector para el que $f_c(\mathbf{v}) < 0$. Comprende, pues, cierto número (inferior a n) de signos $+$.

Observemos, finalmente, que el signo de igualdad en la relación de Schwartz no implica la dependencia: basta considerar dos vectores cualesquiera de los que uno de ellos pertenezca al núcleo.

El caso de formas cuadráticas *definidas negativas* o *semidefinidas negativas* se trata de la misma manera.

Forma indefinida.—Cuando una forma puede tomar valores tanto positivos como negativos se llama *indefinida*. Para ello, es necesario y suficiente que su signatura contenga signos $+$ y $-$; puede ser o no degenerada y no es válida la desigualdad de Schwartz.

EJERCICIOS SOBRE EL CAPITULO XVIII

18-1. Sean x_1, x_2 y x_3 las coordenadas homogéneas de un punto del plano; X designará la matriz columna x_1, x_2, x_3. Sea A una matriz real, simétrica y cuadrada de orden 3, en la que el elemento de la fila i y columna j se denota por a_{ij}. Se considera la cónica (α), cuya ecuación es $\Sigma a_{ij}x_ix_j=0$.

1.º Demuéstrese que la ecuación anterior puede escribirse $X'AX=0$.

2.º ¿A qué condición debe satisfacer A para que (α) sea irreducible?

3.º Suponemos que (α) es irreducible. A un punto M cualquiera del plano se asocia su polar $P_\alpha(M)$ con relación a (α)[1]. Sea (β) otra cónica también irreducible definida de manera análoga por la matriz B, de elementos b_{ij}. Sea M' el polo de $P_\alpha(M)$ con relación a (β).

Se define así una transformación T que hace corresponder a todo punto M un punto M'. Pruébese que T queda definida directamente a partir de las matrices A y B.

¿Es todo punto del plano el transformado por T de un punto M? ¿Cuál es el lugar de M' cuando M describe una recta?

18-2. Demuéstrese que una forma cuadrática 'XAX es descomponible en un producto de dos factores lineales si, y solo si, su rango es inferior o igual a 2.

18-3. Para que una forma cuadrática de tres variables

$$f(x, y, z)=Ax^2+A'y^2+A''z^2+2Byz+2B'zx+2B''xy$$

sea un cuadrado perfecto es necesario y suficiente que se verifiquen las tres relaciones

$$AA'=B''^2, \qquad AA''=B'^2, \qquad AB=B'B''.$$

Considérese el cuerpo de los números complejos. ¿Qué modificación es necesario introducir en el enunciado si se trata del cuerpo de los números reales y se supone que se verifican las tres condiciones anteriores?

---

[1] Obsérvese que la condición para que dos puntos M_1 y M_2 sean conjugados con relación a (α) se escribe $'X_1AX_1=0$.

CAPITULO XIX

ESPACIO EUCLIDIANO. ESPACIO UNITARIO

PRODUCTO ESCALAR

Las teorías desarrolladas hasta aquí reposan únicamente sobre la estructura lineal del espacio estudiado. Nunca hemos hablado de longitud ni de ángulo. Sin embargo, la introducción, en un espacio vectorial definido sobre el cuerpo de los números reales, de una función bilineal y simétrica convenientemente elegida, permite definir estos conceptos de manera que la geometría así obtenida sea, en el caso de dos o tres dimensiones, idéntica a la del espacio ordinario. La teoría que sigue se desarrolla con independencia de los resultados del capítulo anterior [1].

1. **Definición. Propiedad fundamental.**—Se llama *espacio euclidiano* de n dimensiones, sobre el cuerpo R, un espacio vectorial de la misma dimensión en el que está definida una forma bilineal simétrica $f(\mathbf{v}, \mathbf{w})$ tal que la forma cuadrática asociada $f_c(\mathbf{v})$ sea *definida positiva*, es decir, que para todo vector $\mathbf{v}$ se tenga $f_c(\mathbf{v}) \geqslant 0$, siendo la condición $f_c(\mathbf{v}) = 0$ equivalente a $\mathbf{v} = \mathbf{0}$. Esta condición, establecida aquí *a priori*, desempeña un papel capital en lo que sigue [2], y justifica la elección del cuerpo R.

Todo subespacio de un espacio euclidiano es también euclidiano. La función f se llama *producto escalar* [3]; su valor $f(\mathbf{v}, \mathbf{w})$ se representa por la notación clásica $\mathbf{v} \cdot \mathbf{w}$. Para $\mathbf{v} = \mathbf{w}$ se obtiene $f_c(\mathbf{v}) = \mathbf{v} \cdot \mathbf{v}$, que es el *cuadrado escalar* del vector $\mathbf{v}$, representado por $\mathbf{v}^2$; la igualdad $\mathbf{v}^2 = 0$ implica $\mathbf{v} = \mathbf{0}$.

La existencia de un producto escalar en un espacio vectorial entraña una ligazón importante entre este espacio y su dual.

---

[1] Supondremos conocidas, sin embargo, las definiciones fundamentales: bilinealidad, forma cuadrática asociada, escritura matricial, vectores conjugados (págs. 425 a 431).

[2] Se verifica si, y solo si, la signatura de la forma elegida admite n signos + (véase pág. 436).

[3] El producto escalar definido en la página 354 es una función bilineal de vectores que pertenecen a espacios, en general, distintos. Subrayamos la particularidad del caso presente adoptando, para el producto, la notación usual en geometría.

TEOREMA FUNDAMENTAL.—*Todo espacio euclidiano* E *es naturalmente isomorfo a su dual* F *(espacio de las funciones lineales definidas sobre* E).

Sea **a** un vector determinado de E. La función que asocia a todo **v** el escalar $\mathbf{a}\cdot\mathbf{v}$ es lineal en **v**; por consiguiente, es un elemento f_a de F tal que $f_a(\mathbf{v})=\mathbf{a}\cdot\mathbf{v}$. Queda así definida una aplicación de E en F: $\mathbf{a}\longrightarrow f_a$, de la que se demuestra fácilmente que es lineal (bilinealidad del producto y definición de las operaciones sobre F). Hallemos el núcleo: es el conjunto de los $\mathbf{a}\in E$ tales que $f_a=0\in F$; se deduce que

$$\forall\mathbf{v}: \qquad \mathbf{a}\cdot\mathbf{v}=f_a(\mathbf{v})=0.$$

En particular, para $\mathbf{v}=\mathbf{a}$ se tiene $\mathbf{a}^2=0$, de donde $\mathbf{a}=\mathbf{0}$. La aplicación es, pues, inyectiva; los espacios E y F tienen la misma dimensión y, por tanto, es también biyectiva (véase pág. 344). Es un isomorfismo natural, porque se define con independencia de la elección de las bases en E y en F.

Se deduce de ello que a toda función lineal $f\in F$ corresponde un vector determinado $\mathbf{a}\in E$ tal que, para todo **v**, se tenga $f(\mathbf{v})=\mathbf{a}\cdot\mathbf{v}$. Los vectores de E representan biunívocamente las funciones lineales definidas sobre E.

Con frecuencia la correspondencia lineal anterior entre E y F se resume diciendo que un espacio euclidiano es idéntico a su dual.

OBSERVACIÓN.—Hemos razonado directamente, pero el estudio general sobre la dualidad (véase pág. 354) permite encontrar de nuevo fácilmente los resultados anteriores; es preciso solo suponer $E=E^*$ para identificar los dos productos escalares. Además, el carácter definido positivo de la forma $f_a(\mathbf{v})$ implica la condición II de la página 354. El teorema anterior es, pues, un caso particular del teorema 2, página 356.

Algunas de las propiedades que estudiamos aquí (ortogonalidad) se deducen, pues, directamente de la teoría general; sin embargo, las propiedades esenciales adquieren significado porque los dos vectores del producto escalar pertenecen a un mismo espacio y pueden, como caso particular, coincidir. Una base ortonormada, p. ej., es una base que es su propia dual.

2. Forma fundamental.—Se define el producto escalar eligiendo su expresión en una base particular. Según lo dicho en la página 425, esto equivale a precisar los productos escalares de los vectores de la base

tomados dos a dos. Hacemos

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j \qquad \mathbf{g}_{ii} = \mathbf{e}_i^2,$$

de donde se deduce la expresión general

$$\mathbf{v} \cdot \mathbf{w} = \sum g_{\alpha\beta} x^\alpha y^\beta = {}^t\mathrm{XGY},$$

en la que $G = [g_{ij}]$.

Sigue siendo necesario que la forma cuadrática asociada sea definida positiva; en particular, todos los g_{ii} deben ser positivos, aunque esto no es suficiente [1]. Las condiciones sobre la signatura (véase nota 2 de la página 439) no son expresables directamente con ayuda de los coeficientes.

Existe, sin embargo, una elección muy sencilla de la matriz $[g_{ij}]$ que asegura el carácter definido positivo sobre R; esta es

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \qquad (0 \text{ para } i \neq j,\ 1 \text{ para } i = j).$$

La expresión del producto escalar se convierte en

$$\mathbf{v} \cdot \mathbf{w} = \sum_1^n x^\alpha y^\alpha;$$

y la del cuadrado escalar, en

$$\mathbf{v}^2 = \sum_1^n (x^\alpha)^2.$$

Esta es una suma de cuadrados y, en R, la condición $\mathbf{v}^2 = 0$ exige $x^\alpha = 0$ (α de 1 a n), de donde $\mathbf{v} = \mathbf{0}$.

El criterio anterior puede parecer muy particular; demostraremos, sin embargo, que para una elección cualquiera de la forma fundamental, supuesta definida positiva, existen bases llamadas *ortonormadas* en las que el producto escalar se expresa de la manera dicha. La existencia de estas bases resulta, por otra parte, del capítulo anterior. Sin embargo, el método de ortogonalización de Schmidt (pág. 445) permite precisar aquí la construcción de una base ortonormada a partir de una base dada [2].

[1] Así, p. ej., para dos variables $g_{11}(x^1)^2 + 2g_{12}x^1x^2 + g_{22}(x^2)^2$ es definida positiva para $g_{11} > 0$, $(g_{12})^2 - g_{11}g_{22} < 0$ (estrictamente).

[2] Por el contrario, el teorema de la página 431 no precisa la manera de obtener la base $\{\mathbf{e}_1, \mathbf{f}_2, \mathbf{g}_3, \ldots\}$ a partir de una base dada.

3. Ortogonalidad.—Dos vectores $\mathbf{v}$ y $\mathbf{w}$ se dicen *ortogonales* (en lugar de conjugados) cuando su producto escalar es nulo: $\mathbf{v}\cdot\mathbf{w}=0$. Indiquemos algunas consecuencias de esta definición.

a) Un vector ortogonal a sí mismo es nulo: $\mathbf{v}^2=0$ exige $\mathbf{v}=\mathbf{0}$.

b) El vector nulo es ortogonal a todo vector del espacio, y es el único que goza de tal propiedad porque ha de ser, en particular, ortogonal a sí mismo.

c) Un vector es ortogonal a un subespacio si lo es a todo vector del mismo. Para ello, es necesario y suficiente que sea ortogonal a los vectores de una base del subespacio. En particular, un vector ortogonal a los n vectores de una base del espacio es ortogonal a sí mismo; por consiguiente, nulo.

d) Si $\mathbf{v}\neq\mathbf{0}$, un vector cualquiera $\mathbf{w}$ es la suma de un vector proporcional a $\mathbf{v}$ y de un vector ortogonal a $\mathbf{v}$. La descomposición es única.

Este es el teorema 1 de la página 431, válido aquí bajo la sola hipótesis de que sea $\mathbf{v}\neq\mathbf{0}$. Escribimos *a priori* $\mathbf{w}=\lambda\mathbf{v}+\mathbf{w}_1$ y la condición $\mathbf{w}_1\cdot\mathbf{v}=0$ permite calcular λ; se obtiene:

$$\mathbf{w}=\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}^2}\,\mathbf{v}+\mathbf{w}_1.$$

El primer vector del segundo miembro se llama *proyección ortogonal* de $\mathbf{w}$ sobre $\mathbf{v}$.

Desigualdad de Schwartz.—Efectuando el cuadrado escalar de la igualdad anterior se obtiene, puesto que $\mathbf{v}$ y $\mathbf{w}_1$ son ortogonales,

$$\mathbf{w}^2=\frac{(\mathbf{v}\cdot\mathbf{w})^2}{\mathbf{v}^2}+\mathbf{w}_1^2;$$

tenemos $\mathbf{w}_1^2\geqslant 0$ y, por tanto, $(\mathbf{v}\cdot\mathbf{w})^2\leqslant\mathbf{v}^2\cdot\mathbf{w}^2$.

La igualdad exige $\mathbf{w}_1=\mathbf{0}$; por consiguiente, $\mathbf{v}$ y $\mathbf{w}$ dependientes.

Teorema de independencia.—*Todo sistema formado por p vectores no nulos y ortogonales dos a dos es libre.*

Sea el sistema $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ que responde a las condiciones anteriores. Una relación de dependencia

$$\lambda^1\mathbf{v}_1+\lambda^2\mathbf{v}_2+\ldots+\lambda^p\mathbf{v}_p=\mathbf{0}$$

conduce, después de multiplicar escalarmente por $\mathbf{v}_i$ (donde i es un entero determinado) a $\lambda^i\mathbf{v}_i^2=0$, de donde $\lambda^i=0$, puesto que $\mathbf{v}_i\neq\mathbf{0}$. El sistema es, por tanto, libre.

4. Longitudes.—La definición del producto escalar permite dar significado a las nociones de longitud y de ángulo, siendo las propiedades ligadas a estas nociones las mismas que en el espacio ordinario.

Por definición, se llama *longitud* o *norma* de un vector $\mathbf{v}$ la raíz cuadrada positiva de su cuadrado escalar. Se representa por $|\mathbf{v}|$, de donde $|\mathbf{v}| = \sqrt{\mathbf{v}^2}$. Todo vector de longitud nula es, por tanto, nulo. Empleando longitudes, la desigualdad de Schwartz se escribe:

$$-|\mathbf{v}| \cdot |\mathbf{w}| \leqq \mathbf{v} \cdot \mathbf{w} \leqq |\mathbf{v}| \cdot |\mathbf{w}|.$$

Un vector de longitud 1 se dice unitario. La multiplicación de un vector por $\lambda \in \mathbf{R}$ multiplica su longitud por $|\lambda|$ (valor absoluto), ya que $\lambda\mathbf{v} \cdot \lambda\mathbf{v} = \lambda^2\mathbf{v}^2$. Eligiendo $\lambda = \dfrac{1}{\sqrt{\mathbf{v}^2}}$ se obtiene el vector unitario $\dfrac{\mathbf{v}}{\sqrt{\mathbf{v}^2}}$ asociado a $\mathbf{v}$.

Desigualdad de Minkowski.—*La longitud de la suma de dos vectores es inferior o igual a la suma de las longitudes de estos vectores:*

$$|\mathbf{v} + \mathbf{w}| \leqq |\mathbf{v}| + |\mathbf{w}|.$$

Efectuando el cuadrado escalar de $\mathbf{v} + \mathbf{w}$ y utilizando la desigualdad de Schwartz, se obtiene:

$$|\mathbf{v} + \mathbf{w}|^2 = |\mathbf{v}|^2 + 2\mathbf{v} \cdot \mathbf{w} + |\mathbf{w}|^2 \leqq |\mathbf{v}|^2 + 2|\mathbf{v}| \cdot |\mathbf{w}| + |\mathbf{w}|^2$$

y basta extraer la raíz cuadrada.

La igualdad

$$|\mathbf{v} + \mathbf{w}| = |\mathbf{v}| + |\mathbf{w}|$$

solo se verifica si

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}| \cdot |\mathbf{w}|;$$

es decir, para dos vectores proporcionales. Poniendo $\mathbf{w} = \lambda\mathbf{v}$ y excluyendo el caso de los vectores nulos, se obtiene $\lambda = |\lambda|$, o sea que λ debe ser positivo. Geométricamente, los dos vectores han de tener la misma dirección y el mismo sentido.

5. Angulos.—Sean $\mathbf{v}$ y $\mathbf{w}$ dos vectores no nulos. La desigualdad de Schwartz se puede escribir en la forma

$$-1 \leqq \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}| \cdot |\mathbf{w}|} \leqq 1.$$

Existe, pues, un número real, α, del intervalo $(0, \pi)$ tal que

$$\cos \alpha = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}| \cdot |\mathbf{w}|}$$

Este número α se llama ángulo de los vectores $\mathbf{v}$ y $\mathbf{w}$. Es una función simétrica de los dos vectores, igual que el producto escalar. La introducción del ángulo α da al producto escalar su forma clásica: producto de las longitudes por el coseno del ángulo; o sea,

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}| \cdot |\mathbf{w}| \cos \alpha.$$

Si los vectores son ortogonales, $\cos \alpha = 0$, de donde $\alpha = \frac{\pi}{2}$. Si son proporcionales, $\mathbf{w} = \lambda \mathbf{v}$, se tiene que $\cos \alpha = \frac{\lambda}{|\lambda|} = \pm 1$; $+1$ para $\lambda > 0$, $\alpha = 0$, y -1 para $\lambda < 0$, $\alpha = \pi$.

Aplicación.— Consideremos de nuevo la fórmula de la proyección (véase pág. 442)

$$\mathbf{w} = \lambda \mathbf{v} + \mathbf{w}_1 \qquad \text{con} \qquad \lambda = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|^2} = \frac{|\mathbf{w}|}{|\mathbf{v}|} \cos \alpha,$$

donde α es el ángulo de $\mathbf{v}$ y $\mathbf{w}$. Si $\mathbf{v}_1$ designa el vector unitario asociado a $\mathbf{v}$, la proyección de $\mathbf{w}$ sobre $\mathbf{v}$ se escribe $|\mathbf{w}| \cos \alpha \, \mathbf{v}_1$. El escalar $|\mathbf{w}| \cos \alpha$ es el valor algebraico de la proyección. Se vuelve así a encontrar el teorema de la proyección de la geometría métrica corriente.

Finalmente, si $\mathbf{w}$ es unitario, la proyección de $\mathbf{w}$ tiene por valor algebraico $\cos \alpha$; esta es la definición clásica del coseno de un ángulo.

Relación del coseno.—En la relación

$$|\mathbf{v} + \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 + 2\mathbf{v} \cdot \mathbf{w}$$

expresemos el producto escalar con ayuda del coseno; se obtiene:

$$|\mathbf{v} + \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 + 2|\mathbf{v}| \cdot |\mathbf{w}| \cos \alpha \qquad \alpha = (\mathbf{v}, \mathbf{w}).$$

Cambiando $\mathbf{w}$ en $-\mathbf{w}$, resulta:

$$|\mathbf{v} - \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 - 2|\mathbf{v}| \cdot |\mathbf{w}| \cos \alpha,$$

que es, en la geometría de n dimensiones, la relación clásica del triángulo.

BASES ORTONORMADAS. PRODUCTO VECTORIAL

6. Método de ortogonalización de Schmidt. TEOREMA.—*En un espacio euclidiano de n dimensiones todo subespacio de p dimensiones puede engendrarse por un sistema de p vectores unitarios y ortogonales dos a dos.*

Sea $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_p$ una base del subespacio considerado, y hallemos en primer lugar una base constituida por vectores $\mathbf{f}'_1, \mathbf{f}'_2, \ldots, \mathbf{f}'_p$ ortogonales dos a dos. El método de ortogonalización de Schmidt generaliza la descomposición de un vector estudiada en la página 442. Se elige $\mathbf{f}'_1 = \mathbf{f}_1$; hemos visto que $\mathbf{f}_2$ puede escribirse $\mathbf{f}_2 = \lambda \mathbf{f}'_1 + \mathbf{f}'_2$ con $\mathbf{f}'_2$ ortogonal a $\mathbf{f}'_1$ y no nulo en este caso, ya que $\mathbf{f}'_1$ y $\mathbf{f}_2$ son independientes. Se determinan a continuación λ' y μ' tales que

$$\mathbf{f}'_3 = \mathbf{f}_3 - \lambda' \mathbf{f}'_1 - \mu' \mathbf{f}'_2$$

sea ortogonal a $\mathbf{f}'_1$ y $\mathbf{f}'_2$. Tenemos:

$$0 = \mathbf{f}_3 \cdot \mathbf{f}'_1 - \lambda' |\mathbf{f}'_1|^2 \qquad \text{y} \qquad 0 = \mathbf{f}_3 \cdot \mathbf{f}'_2 - \mu' |\mathbf{f}'_2|^2.$$

Estas dos igualdades determinan de manera única λ' y μ'. El vector $\mathbf{f}'_3$ así obtenido es ciertamente no nulo, porque $\mathbf{f}_3$, $\mathbf{f}'_2$ y $\mathbf{f}'_1$, o también $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$, son independientes. El proceso se prosigue descomponiendo $\mathbf{f}_4$, etcétera.

Geométricamente, el método consiste en descomponer todo vector $\mathbf{f}_i$ en una componente en el subespacio engendrado por $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{i-1}$ y otra ortogonal a dicho subespacio. Cualquiera que sea i, es posible esta descomposición y la componente ortogonal no es nula, puesto que $\mathbf{f}_i$ no pertenece al subespacio engendrado por $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{i-1}$.

De esta manera, se obtendrá para el subespacio de p dimensiones una base $\mathbf{f}'_1, \mathbf{f}'_2, \ldots, \mathbf{f}'_p$ formada por vectores ortogonales dos a dos. Basta entonces normalizar estos vectores; es decir, formar la base

$$\frac{\mathbf{f}'_1}{\sqrt{|\mathbf{f}'_1|^2}}, \frac{\mathbf{f}'_2}{\sqrt{|\mathbf{f}'_2|^2}}, \ldots, \frac{\mathbf{f}'_p}{\sqrt{|\mathbf{f}'_p|^2}},$$

que se llama *ortonormada.*

Base ortonormada de E.—El procedimiento anterior puede aplicarse a todo el espacio E. A partir de cualquier base se obtiene otra *base or-*

tonormada $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ tal que $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. Ya que la matriz de la forma fundamental es I, el producto escalar se escribe

$$\mathbf{v} \cdot \mathbf{w} = {}^t\mathrm{X}\mathrm{Y} = \mathrm{Y}^t\mathrm{X} = \sum_1^n x^\alpha y^\alpha.$$

En toda base ortonormada la desigualdad de Schwartz se traduce por la relación [1]:

$$(\Sigma x^\alpha y^\alpha)^2 \leqslant \Sigma(x^\alpha)^2 \cdot \Sigma(y^\alpha)^2.$$

Finalmente, si $\mathbf{v} = \Sigma x^\alpha \mathbf{e}_\alpha$ es un vector cualquiera, se tiene:

$$\mathbf{v} \cdot \mathbf{e}_i = x^i = |\mathbf{v}| \cos(\mathbf{v}, \mathbf{e}_i).$$

En una base ortonormada las coordenadas de un vector son los productos escalares de este con los vectores de la base. Si v es unitario, los n números $x^i = \cos(\mathbf{v}, \mathbf{e}_i)$, tales que $\Sigma(x^\alpha)^2 = 1$, se llaman *cosenos directores* de v.

Consecuencias.—Queda, pues, establecido el resultado enunciado en la página 442. La elección arbitraria de una forma cuadrática definida positiva equivale a elegir una base considerada como ortonormada, y recíprocamente.

Finalmente, el isomorfismo entre dos espacios vectoriales reales E y E′ de la misma dimensión se extiende a las estructuras euclidianas, si se ha obtenido asociando dos bases ortonormadas, una de E y la otra de E′. Si a $\mathbf{v} \in \mathrm{E}$ corresponde $\mathbf{v}' \in \mathrm{E}'$, y a w, w′, tenemos, en efecto,

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}' \cdot \mathbf{w}'.$$

En particular, todo espacio euclidiano real de n dimensiones es isomorfo con el espacio vectorial R^n, en el que la base canónica se considera como ortonormada. Lo llamamos *espacio euclidiano* R^n, y el isomorfismo anterior justifica su importancia.

---

[1] Llamada también desigualdad de Cauchy-Schwartz. Se obtiene también por cálculo escribiendo que el trinomio en λ,

$$\varphi(\lambda) = \sum_1^n (\lambda x^\alpha + y^\alpha)^2$$

es definido positivo. Es válida para dos sistemas cualesquiera de n números reales.

7. Suplemento ortogonal.—Sean dos subespacios: E′ engendrado por $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ y E″ engendrado por $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q\}$. Si suponemos $\mathbf{v}_i \cdot \mathbf{w}_j = 0$, cualesquiera que sean i (de 1 a p) y j (de 1 a q), es fácil ver que la ortogonalidad se amplía a las combinaciones lineales; es decir, que todo vector de E′ es ortogonal a todo vector de E″. Se dice que E′ y E″ son dos subespacios ortogonales. El siguiente teorema considera el caso en el que estos dos subespacios son, además, suplementarios.

TEOREMA.—*En un espacio euclidiano de n dimensiones todo subespacio de p dimensiones admite un subespacio ortogonal suplementario, y uno solo.*

Cabe suponer el subespacio E'_p referido a una base ortonormada $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p\}$, y sea $\{\mathbf{f}_{p+1}, \mathbf{f}_{p+2}, \ldots, \mathbf{f}_n\}$ una base de un subespacio suplementario. Apliquemos el método de ortogonalización de Schmidt a la base $\{\mathbf{e}_1, \ldots, \mathbf{e}_p, \mathbf{f}_{p+1}, \ldots, \mathbf{f}_n\}$ de E así formada. Como los p primeros vectores no cambian, se obtiene una base ortonormada $\{\mathbf{e}_1, \ldots, \mathbf{e}_p, \mathbf{e}_{p+1}, \ldots, \mathbf{e}_n\}$ tal que el subespacio E'_{n-p} engendrado por los $n-p$ últimos vectores es ortogonal al subespacio engendrado por los p primeros; dicho subespacio es también suplementario.

Demostremos que es único. Si hubiera otro, engendrado por $\{\mathbf{h}_{p+1}, \ldots, \mathbf{h}_n\}$, para todo $\mathbf{h}_i$ se tendría:

$$\mathbf{h}_i = \sum_{1}^{p} \lambda^\alpha \mathbf{e}_\alpha + \sum_{p+1}^{n} \lambda^\alpha \mathbf{e}_\alpha.$$

Los productos $\mathbf{h}_a \cdot \mathbf{e}_\alpha$ (α de 1 a p) deberían ser nulos y, por consiguiente, todos los coeficientes λ^α (α de 1 a p) serían ceros. Puesto que lo mismo sucedería para todos los vectores $\mathbf{h}$, los dos subespacios coincidirían.

EJERCICIOS:

1. En el espacio euclidiano R^4 determínese un vector unitario ortogonal a los tres vectores (3, 2, 2, 4), (1, 0, 0, 2), (1, −1, −1, 1).

2. Determínese en R^4 una base del subespacio que tiene por ecuaciones

$$x^1 + 3x^2 - x^3 + 2x^4 = 0, \qquad 2x^1 - x^2 + 3x^3 - x^4 = 0.$$

¿Cuál es el subespacio ortogonal y suplementario?

3. En el espacio euclidiano R^4 descompóngase $\mathbf{v} = (2, 3, -1, 4)$ en suma de dos vectores, uno perteneciente al subespacio engendrado por $\mathbf{a}_1 = (2, 1, -1, 1)$ y $\mathbf{a}_2 = (3, 1, 1, 0)$, y el otro ortogonal a este subespacio.

4. Se considera en R^4 un subespacio de dos dimensiones E_1 y un vector w cualquiera. Defínase la proyección ortogonal w' de w sobre E_1 y compruébese que w' es, entre los vectores de E_1, el que forma menor ángulo con w. ¿Existe algún caso de excepción?

Ejemplo.—Sea el subespacio engendrado por $a_1=(-1, 1, -1, 1)$ y $a_2=(1, 5, -3, 3)$. Hállese el vector del subespacio que forma ángulo mínimo con $v=(3, 1, 1, -3)$.

5. Sean E_1 y E_2 los dos subespacios de R^4, respectivamente engendrados por (1, 0, 0, 0), (0, 1, 0, 0) y (1, 1, 1, 1), (2, −2, 5, 2). A todo par de vectores $v \in E_1$ y $v' \in E_2$ asociamos el ángulo (v, v'). Determínese el par al que corresponde el ángulo mínimo.

Indicación.—Evidentemente, el vector v' buscado es el vector de E_2 que forma el ángulo mínimo con su proyección sobre E_1.

6. Hállense en el espacio euclidiano R^5 dos vectores ortogonales entre sí y ortogonales al subespacio engendrado por (1, 0, 0, 1, −2), (2, 1, −1, 0, 2), (1, 1, 1, 2, 1).

Entre los pares posibles precísense aquellos para los que uno de los vectores tiene nula su primera coordenada.

7. Sean E_1, E_2, ..., E_p, p subespacios ortogonales dos a dos de un espacio de n dimensiones. Demuéstrese que la suma $E_1+E_2+\ldots+E_p$ es directa.

8. Sea $E=E_1 \oplus E_2$ una descomposición de E en suma directa. Demuéstrese que los suplementos ortogonales E_1' y E_2' de E_1 y E_2 constituyen una descomposición de E en suma directa.

9. Demuéstrese que en un espacio euclidiano de n dimensiones es posible determinar n vectores unitarios $u_1, u_2, \ldots, u_n$ tales que, para todo par de enteros i y j, el vector u_i-u_j sea igualmente unitario. (Razónese por recurrencia.) ¿Es posible determinar $n+1$ vectores que verifiquen las mismas condiciones?

Hacemos

$$v=-\frac{1}{n+1}\sum_{\alpha} u_\alpha \quad (\alpha \text{ de } 1 \text{ a } n).$$

Calcúlese el ángulo de los dos vectores u_i y u_j, así como el ángulo de los vectores $v-u_i$ y $v-u_j$.

8. Cambio de base. Matrices ortogonales.—Por lo general, en geometría euclidiana de n dimensiones se eligen bases ortonormadas. Vamos a caracterizar las matrices de cambio P que asocian tales bases. En general, las matrices G y G_1 de la forma fundamental son tales que

$$G_1={}^tPGP;$$

pero aquí $G=G_1=I$, de donde

$${}^tPP=I. \qquad [1]$$

La traspuesta de P es también su inversa: $P^{-1} = {}^tP$, lo que implica, además, $P\,{}^tP = I$. Finalmente, esta condición es suficiente, porque con $G = I$ se obtiene:

$$G_1 = {}^tPIP = {}^tPP = I.$$

TEOREMA.—*Una matriz de cambio* P *asocia dos bases ortonormadas si, y solo si, su inversa es igual a su traspuesta.*

Las matrices que satisfacen esta condición se llaman *ortogonales*.

Condiciones de ortogonalidad.—Se escriben fácilmente las condiciones de ortogonalidad de una matriz

$$P = \begin{bmatrix} \alpha_1^1 & \alpha_2^1 & \dots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \vdots & & & \\ \alpha_1^n & \alpha_2^n & \dots & \alpha_n^n \end{bmatrix}$$

en la que, siguiendo la costumbre, ordenamos en una columna las componentes de uno de los nuevos vectores de la base. La relación ${}^tPP = I$ se traduce por las n^2 relaciones escalares

$$\sum_{k=1}^{k=n} \alpha_i^k \alpha_j^k = \delta_{ij}, \qquad [2]$$

que expresan geométricamente que los n nuevos vectores de la base son unitarios y ortogonales dos a dos.

Puesto que ${}^tP = P^{-1}$, en las filas de P se leen los vectores columna de P^{-1}, es decir, las coordenadas de los antiguos vectores de la base en la nueva base. Aparecen así n^2 nuevas relaciones

$$\sum_{k=1}^{k=n} \alpha_k^i \alpha_k^j = \delta^{ij}, \qquad [3]$$

que expresan, en la nueva base, la ortogonalidad de la antigua. Naturalmente, uno cualquiera de estos sistemas de relaciones es consecuencia del otro, porque cada uno de ellos expresa la ortogonalidad de P. Esta es la generalización de las relaciones clásicas que caracterizan un cambio de ejes rectangulares en la geometría de tres dimensiones.

Determinante.—El determinante de una matriz ortogonal P es igual a ± 1. En efecto, tenemos según [1],

$$|{}^tP| \cdot |P| = |I| \qquad \text{o} \qquad |P|^2 = 1 \qquad \text{y} \qquad |P| = \pm 1.$$

La existencia de dos clases de matrices ortogonales se relaciona con la noción de orientación que precisamos a continuación. Las matrices tales que $|P| = +1$ se llaman ortogonales directas; aquellas para las cuales $|P| = -1$, ortogonales inversas.

9. Función determinante en un espacio euclidiano. Producto vectorial.—Hemos visto (véase pág. 373) que una función determinante se caracteriza por la elección de su valor para una base del espacio. El siguiente teorema precisa el caso de las bases ortonormadas.

TEOREMA.—*En un espacio euclidiano sobre* R, *toda función determinante toma, salvo el signo, el mismo valor para cualquier base ortonormada.*

Si P es la matriz de cambio de la base $\{\mathbf{e}_i\}$ a la base $\{\mathbf{f}_i\}$, se tiene para una función determinante cualquiera Δ:

$$\Delta(\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n) = |P| \Delta(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n).$$

Pero sabemos que $|P| = \pm 1$, de donde resulta el teorema.

Consecuencias.—*a)* Por consiguiente, en un espacio euclidiano resulta natural normalizar una función determinante precisando que toma el valor I para una base ortonormada particular. Tomará entonces el valor ± 1 para cualquier otra base ortonormada.

b) Esta elección permite también orientar el espacio (véase página 374). Las bases ortonormadas positivas corresponden al valor $+1$ de la función considerada, y las negativas, al valor -1.

Se ve igualmente que dos bases de la misma orientación están ligadas por una matriz de cambio ortogonal directa, mientras que dos bases de orientaciones diferentes lo están por una matriz ortogonal inversa.

Producto vectorial.—Llamemos Δ la función determinante a que nos hemos referido (valor $+1$ para toda base positiva). Consideremos un sistema de $n-1$ vectores dados $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$. Si $\mathbf{w}$ es un vector arbitrario, la aplicación de E en R,

$$\mathbf{w} \longrightarrow \Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}, \mathbf{w})$$

es lineal con relación a $\mathbf{w}$. Existe, pues, un vector bien determinado $\mathbf{a} \in E$ tal que

$$(\forall \mathbf{w}) \quad \Delta(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}, \mathbf{w}) = \mathbf{a} \cdot \mathbf{w}. \qquad [1]$$

Por definición, este vector $\mathbf{a}$ es el producto vectorial de los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. Esta es la particularización al caso $E = E^*$ de la definición de producto exterior (véase pág. 391). El lector precisará para este caso concreto las propiedades de la operación. Las recordamos a continuación.

1) Es lineal y antisimétrica con relación a los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$.

2) La condición $\mathbf{a} = 0$ equivale a la dependencia del sistema $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$.

3) La fórmula [1] prueba que en una base ortonormada positiva, las coordenadas de $\mathbf{a}$ son los cofactores de los elementos de la *n-ésima* columna de un determinante del que las $n-1$ primeras columnas están formadas por las coordenadas de $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$.

EJERCICIOS:

1. Demuéstrese que para el orden 2 toda matriz ortogonal directa es de la siguiente forma $\begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$, y toda matriz ortogonal inversa, de la forma $\begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{bmatrix}$.

2. Si $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ es una base ortonormada de R^3 demuéstrese que sucede lo mismo con la base formada por los vectores $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ de coordenadas respectivas (2/3, 2/3, −1/3), (2/3, −1/3, 2/3), (1/3, −2/3, −2/3). Precísese la orientación relativa de estas dos bases.

3. En R^3, referido a una base ortonormada, se consideran los cambios de base que corresponden, respectivamente, a las matrices de cambio

$$P = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{bmatrix}, \qquad P_1 = \frac{1}{7}\begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix}$$

Compruébese que las nuevas bases son ortonormadas. ¿Cuál es la matriz que efectúa el cambio de la segunda base a la tercera?

4. Determínese una base ortonormada de R^4 de la que los dos primeros vectores tienen por componentes (1/2, 1/2, 1/2, 1/2) y (1/6, 1/6, 1/2, −5/6) en una base ortonormada dada, y tal que el tercer vector tenga su primera componente nula.

OPERADORES LINEALES: TRASPUESTO, OPERADOR ORTOGONAL

La introducción de un producto escalar en un espacio vectorial permite asociar a todo operador lineal un operador *traspuesto*, lo que particulariza en un caso sencillo[1] e importante el estudio hecho en el capítulo 18 (pág. 426). Estudiamos a continuación ciertos operadores en relación sencilla con sus traspuestos.

10. Traspuesto de un operador.—Consideremos un espacio euclidiano E de n dimensiones sobre R; sea T un endomorfismo de E, y $\mathbf{a}$ un vector dado. Si $\mathbf{v}$ es un vector arbitrario de E, la aplicación (de E en R) que asocia a $\mathbf{v}$ el escalar $\mathbf{a}\cdot T\mathbf{v}$ es lineal con relación a $\mathbf{v}$ (linealidad del producto escalar y del operador T). En virtud del teorema fundamental (pág. 440) existe un vector $\mathbf{a}' \in E$ tal que

$$\forall \mathbf{v}: \quad \mathbf{a}\cdot T\mathbf{v} = \mathbf{a}'\cdot \mathbf{v}.$$

La igualdad anterior, válida para todo $\mathbf{a} \in E$, define una aplicación, $\mathbf{a} \longrightarrow \mathbf{a}'$, de E en sí mismo. Escribimos $\mathbf{a}' = {}^tT\mathbf{a}$. La bilinealidad del producto escalar implica, como se demuestra fácilmente, la linealidad de ${}^tT$; este es, por definición, el *operador traspuesto*.

TEOREMA.—*A todo operador lineal* T *de un espacio euclidiano se puede asociar un operador traspuesto,* ${}^tT$, *tal que para* $\mathbf{v}$ *y* $\mathbf{w}$ *arbitrarios se tenga*

$$\mathbf{v}\cdot T\mathbf{w} = \mathbf{w}\cdot {}^tT\mathbf{v}.$$

La operación es recíproca; en efecto, el traspuesto de ${}^tT$ es T' tal que

$$\mathbf{w}\cdot {}^tT\mathbf{v} = \mathbf{v}\cdot T'\mathbf{w} = \mathbf{v}\cdot T\mathbf{w}.$$

La última igualdad, que se verifica para $\mathbf{v}$ arbitrario, exige $T'\mathbf{w} = T\mathbf{w}$, de donde $T' = T$ puesto que $\mathbf{w}$ es también arbitrario.

Núcleo.—Sea $\mathbf{w}$ un vector que pertenece al núcleo de T. Puesto que $T\mathbf{w} = \mathbf{0}$, se tiene:

$$\forall \mathbf{v}: \quad \mathbf{v}\cdot T\mathbf{w} = \mathbf{w}\cdot {}^tT\mathbf{v} = 0.$$

---

[1] El estudio general utiliza cuatro espacios vectoriales. Aquí están confundidos en uno solo, lo que justifica su estudio directo.

Por consiguiente, w es ortogonal a todo vector ^tTv; es decir, al subespacio imagen de ^tT. Recíprocamente, sea w un vector que goce de tal propiedad; para v arbitrario,

$$0 = w \cdot {}^tTv = v \cdot Tw.$$

En consecuencia, Tw, ortogonal a todo $v \in E$, es nulo; w pertenece al núcleo de T.

TEOREMA.—*El núcleo del operador* T *(respectivamente,* ^tT*) es el subespacio ortogonal y suplementario del subespacio imagen de* ^tT *(respectivamente,* T*).*

Si r es el rango de ^tT, el núcleo de T tiene por dimensión $n - r$; según el teorema de la página 343, el subespacio imagen de T posee también dimensión r. Los dos operadores T y ^tT tienen el mismo rango.

Matrices.—Sean A y B las matrices de T y ^tT en una base ortonormada de E, y X e Y, las matrices columna de v y w. La igualdad que define la trasposición se escribe:

$$^tXAY = {}^tYBX.$$

Por ser escalares ambas expresiones, podemos sustituir una de ellas por su traspuesta, de donde

$$^tXAY = {}^tX{}^tBY, \qquad \text{o sea} \qquad {}^tX(A - {}^tB)Y = 0.$$

Esta igualdad, válida para X e Y arbitrarios, implica $A = {}^tB$. Recíprocamente, esta condición prueba que los dos operadores son traspuestos.

TEOREMA.—*Dos operadores de un espacio euclidiano son traspuestos uno del otro si, y solo si, admiten dos matrices traspuestas una de la otra en una misma base ortonormada.*

Valores propios.—La condición anterior sobre las matrices demuestra que dos operadores traspuestos admiten la misma ecuación característica y tienen, por consiguiente, los mismos valores propios (suponiendo que estos existan en R). Las trazas de los dos operadores y sus determinantes son iguales. Tenemos, finalmente, el teorema siguiente.

TEOREMA.—*Dos subespacios propios, uno de* T *y otro de* ^tT*, que corresponden a dos valores propios distintos son ortogonales.*

Sean v tal que $Tv=\lambda v$ y w tal que ${}^tTw=\lambda_1 w$. La relación

$$w\cdot Tv=v\cdot {}^tTw$$

se escribe

$$\lambda w\cdot v=\lambda_1 v\cdot w,$$

o

$$(\lambda-\lambda_1)v\cdot w=0.$$

Puesto que $\lambda-\lambda_1\neq 0$, los dos vectores son ortogonales. La propiedad se extiende inmediatamente a los dos subespacios.

EJERCICIOS:

1. Sean T y T' dos operadores lineales de un espacio euclidiano. Demuéstrense las fórmulas

$${}^t(T+T')={}^tT+{}^tT'; \qquad {}^t(T'\circ T)={}^tT\circ {}^tT'.$$

2. Referido el espacio euclidiano a una base cualquiera, sea G la matriz de la forma fundamental. Indíquese la relación entre las matrices A y B de dos operadores traspuestos. Compruébese que esta relación se conserva en todo cambio de base.

3. En R^3 referido a una base ortonormada $\{e_1, e_2, e_3\}$ se considera la proyección paralela al vector (λ, μ, ν) sobre el subespacio $\{e_1, e_2\}$. Estúdiese el operador traspuesto.

4. Sea T un operador de proyección (véase pág. 345, ejercicio 2) del espacio euclidiano R^n. Demuéstrese que el operador traspuesto ${}^tT$ es también una proyección. Caracterícense el subespacio imagen y el núcleo de ${}^tT$.

5. Estúdiese el operador T que en una base ortonormada de R^3 admite la matriz

$$A=\begin{bmatrix} -a & a & 0 \\ 0 & -a & a \\ a & 0 & -a \end{bmatrix}.$$

Determínense el núcleo y el subespacio imagen. Caracterícese geométricamente T eligiendo para ello una base conveniente de R^3.

6. Sea T(E) el subespacio imagen del operador T. Si $w\in T(E)$, demuéstrese que w admite un original único v, ortogonal al núcleo de T.

11. Transformación ortogonal. Grupo ortogonal.—En el espacio euclidiano E estudiemos en particular las transformaciones lineales que conservan el producto escalar. Una transformación R de este tipo se llama *ortogonal*. Para v y w arbitrarios, se tiene:

$$v\cdot w=Rv\cdot Rw.$$

Indiquemos algunas consecuencias.

1.ª Toda transformación ortogonal conserva la longitud de un vector cualquiera.

2.ª Recíprocamente, todo operador lineal que conserva la longitud de un vector cualquiera es ortogonal. En efecto, la relación

$$(\mathbf{v}+\mathbf{w})^2=(\mathbf{R}\mathbf{v}+\mathbf{R}\mathbf{w})^2$$

se escribe:

$$\mathbf{v}^2+\mathbf{w}^2+2\mathbf{v}\cdot\mathbf{w}=(\mathbf{R}\mathbf{v})^2+(\mathbf{R}\mathbf{w})^2+2\mathbf{R}\mathbf{v}\cdot\mathbf{R}\mathbf{w}.$$

Puesto que $\mathbf{v}^2=(\mathbf{R}\mathbf{v})^2$, ..., queda $\mathbf{v}\cdot\mathbf{w}=\mathbf{R}\mathbf{v}\cdot\mathbf{R}\mathbf{w}$.

3.ª Una condición necesaria y suficiente para que una transformación sea ortogonal es que transforme una base ortonormada en otra base ortonormada.

Evidentemente, la condición es necesaria porque se conservan las longitudes de los vectores de la base y sus productos escalares. Es también suficiente, dado que si la base ortonormada $\{\mathbf{e}_i\}$ se transforma en $\{\mathbf{f}_i=\mathbf{R}\mathbf{e}_i\}$, el vector $\mathbf{v}=\Sigma x^a\mathbf{e}_a$ se convierte en $\mathbf{R}\mathbf{v}=\Sigma x^a\mathbf{f}_a$; se conservan las longitudes y los productos escalares.

Un cambio de base ortonormada define, pues, la transformación ortogonal R que a todo v asocia el vector Rv cuyas componentes, en la nueva base, son las de v en la antigua.

TEOREMA 1.—*El conjunto de las transformaciones ortogonales forma un grupo respecto del producto de las transformaciones.*

El producto de dos transformaciones ortogonales es, evidentemente, ortogonal y toda transformación ortogonal R admite una inversa; en efecto, estudiemos su núcleo: si $\mathbf{R}\mathbf{v}=\mathbf{0}$ tenemos $(\mathbf{R}\mathbf{v})^2=\mathbf{v}^2=\mathbf{0}$, de donde $\mathbf{v}=\mathbf{0}$. Finalmente, la inversa, que conserva las longitudes, es también ortogonal.

TEOREMA 2.—*Una condición necesaria y suficiente para que una transformación sea ortogonal es que su inversa sea idéntica a su traspuesta.*

Sea R ortogonal, de inversa $\mathbf{R}^{-1}$. Puesto que $\mathbf{R}^{-1}$ conserva el producto escalar, para dos vectores arbitrarios v y w, se tiene:

$$\mathbf{v}\cdot\mathbf{R}\mathbf{w}=\mathbf{R}^{-1}\mathbf{v}\cdot\mathbf{R}^{-1}\mathbf{R}\mathbf{w}=\mathbf{R}^{-1}\mathbf{v}\cdot\mathbf{w}.$$

La igualdad así obtenida demuestra que la traspuesta de R es $\mathbf{R}^{-1}$. En una base ortonormada la matriz de una transformación ortogonal es ortogonal.

Consecuencias.— La relación $R'R = I$ prueba que el determinante de una transformación ortogonal (véase pág. 406) es igual a ± 1. Las transformaciones ortogonales de determinante $+1$ conservan la orientación de una base. Constituyen el grupo de las rotaciones, que es un subgrupo del grupo ortogonal.

Por el contrario, las transformaciones ortogonales de determinante -1 cambian la orientación de una base cualquiera y no constituyen un subgrupo. Son las transformaciones ortogonales inversas.

EJERCICIOS:

1. Pruébese que en el cuerpo de los reales una transformación ortogonal no puede admitir más que los valores propios ± 1.

2. Demuéstrese que si la dimensión n de E es un número impar, toda rotación admite el valor propio $+1$. (Estúdiese el polinomio característico $\Delta(\lambda)$ para λ infinito y para $\lambda = 0$.) Existe, pues, un subespacio invariante de dimensión igual o mayor que uno.

En la misma hipótesis, pruébese que toda transformación ortogonal inversa admite el valor propio -1.

3. Demuéstrese que si la dimensión n es par toda transformación ortogonal inversa admite necesariamente los valores propios ± 1.

4. Pruébese que una transformación ortogonal es un operador simétrico (véase a continuación) si, y solo si, su cuadrado es la transformación idéntica.

OPERADORES Y FORMAS SIMETRICAS

12. Operador simétrico.—Un operador lineal idéntico a su traspuesto se llama simétrico: $T = {}^tT$. Siendo v y w arbitrarios, tenemos:

$$v \cdot Tw = w \cdot Tv.$$

En toda base ortonormada la matriz de tal operador es simétrica, y recíprocamente.

Observemos que la noción de operador simétrico queda condicionada a la estructura euclidiana del espacio y está relacionada con una elección determinada del producto escalar.

Precisemos las propiedades de estos operadores. Ya sabemos (véase pág. 453) que dos vectores propios que corresponden a valores propios distintos son ortogonales.

TEOREMA 1.—*Todo operador simétrico T de un espacio euclidiano de n dimensiones admite n valores propios reales, distintos o confundidos.*

La demostración que sigue utiliza el cuerpo de los complejos; por tanto, no debe buscarse interpretación de la misma en el espacio euclidiano [1].

Sean $A=[a_{ij}]$ la matriz de T en una base ortonormada y λ un valor propio, *a priori* complejo, de T. Le corresponde al menos una matriz columna (compleja) $X=[x^i]$ tal que

$$AX=\lambda X. \qquad [1]$$

Sea $\overline{X}=[\bar{x}^i]$ la conjugada de X; de [1] se deduce la igualdad escalar

$${}^t\overline{X}AX=\lambda\,{}^t\overline{X}X.$$

Comprobemos que los dos miembros, ${}^t\overline{X}X$ y ${}^t\overline{X}AX$, son reales. El primero se escribe

$${}^t\overline{X}X=\sum_1^n \bar{x}^i x^i=\sum_1^n |x^i|^2,$$

y es real positivo. El segundo,

$$\sum_{i,j} a_{ij}x^j\bar{x}^i$$

tiene por conjugado

$$\sum_{ij} a_{ij}\bar{x}^j x^i.$$

Estos números son iguales porque, debido a la simetría ($a_{ij}=a_{ji}$) figuran exactamente los mismos términos en ambas sumas (se han cambiado los índices mudos). De ello resulta que λ, cociente de dos números reales, es real.

Consecuencia.—*Todo operador simétrico posee al menos un vector propio.*

La ecuación característica, de grado n, admite n raíces reales distintas o confundidas. Existe, pues, un valor propio al menos al que corresponde un subespacio propio cuya dimensión es igual o mayor que uno.

[1] Por el contrario, en el espacio unitario este cálculo adquiere toda su significación (véase pág. 471).

TEOREMA 2.—*El subespacio ortogonal a un vector propio de un operador simétrico es invariante para este operador.*

Sean v un vector propio, $Tv = \lambda v$ y v_1 un vector ortogonal. La relación fundamental se escribe:

$$v \cdot Tv_1 = v_1 \cdot Tv = \lambda v_1 \cdot v = 0.$$

Por consiguiente, Tv_1 es ortogonal a v y pertenece, así como v_1, al subespacio ortogonal a v. Este subespacio es invariante respecto a T, propiedad que es fundamental para la diagonalización.

13. Diagonalización de un operador simétrico.—Sea T un operador simétrico. Su polinomio característico se descompone por completo en el cuerpo de los reales. Si las n raíces son distintas puede ser diagonalizado y, además, los vectores propios son ortogonales dos a dos. El siguiente razonamiento demuestra que es posible la diagonalización incluso en el caso de raíces múltiples.

Se ha visto que T admite al menos un vector propio e_1. Ya que el subespacio ortogonal a e_1 es invariante, T induce (véase pág. 410) en este subespacio de $n-1$ dimensiones un operador T_1 que es igualmente simétrico (restricción a este subespacio de la relación $v \cdot Tv_1 = v_1 \cdot Tv$). El operador T_1 admite, pues, al menos un vector propio e_2 ortogonal a e_1; evidentemente, e_2 es vector propio de T.

Basta continuar de este modo. El subespacio ortogonal a e_1 y e_2 es invariante por T, el cual induce en él un operador simétrico, T_2, que tiene al menos un vector propio, e_3, ortogonal a e_1 y e_2. Después de $n-1$ operaciones como máximo habremos obtenido una base ortogonal $e_1, e_2, \ldots, e_n$ que es suficiente normalizar y que está formada por vectores propios. El operador está diagonalizado en esta base.

TEOREMA.—*Existe al menos una base ortonormada de un espacio euclidiano en la que un operador simétrico admite una matriz diagonal. Los elementos diagonales son los valores propios.*

Subespacios propios.—Sean $\lambda_1, \lambda_2, \ldots, \lambda_p$ los valores propios distintos, de órdenes respectivos $k_1, k_2, \ldots, k_p$. Por ser posible la diagonalización sabemos que a la raíz propia λ_i le corresponde un subespacio propio E_i de dimensión k_i. Además, los subespacios propios son ortogonales dos a dos. Su suma directa se confunde con E.

Si la ecuación característica posee raíces múltiples, existe una infinidad de bases ortonormadas que aseguran la diagonalización. Basta para ello elegir arbitrariamente una base ortonormada en cada uno de

los subespacios E_i (i de 1 a p). La asociación de estas p bases constituye una base de E que responde a la cuestión.

El resultado anterior puede expresarse en forma matricial: se puede hacer corresponder a toda matriz real y simétrica A una matriz ortogonal P, al menos, tal que $P^{-1}AP = {}^tPAP$ sea diagonal. De ello resulta que toda matriz simétrica es semejante a una matriz diagonal.

EJEMPLO.—*Determínese una matriz ortogonal que permita diagonalizar la matriz*

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

El polinomio característico admite una raíz doble $\lambda_1 = -1$ y una raíz simple $\lambda_2 = 8$. El subespacio propio E_1 correspondiente a λ_1 está engendrado por los vectores $(1, -2, 0)$ y $(1, 0, -1)$; una base ortonormada es

$$\left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right), \quad \left(\frac{1}{3\sqrt{2}}, \frac{-4}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}\right);$$

a $\lambda_2 = 8$ le corresponde el vector propio normalizado (2/3, 1/3, 2/3). Por consiguiente, una matriz ortogonal que asegura la diagonalización es

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{4}{3\sqrt{2}} & \frac{1}{3} \\ -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \end{bmatrix}$$

La matriz A es semejante a diag. $\{-1, -1, 8\}$.

EJERCICIOS:

1. Demuéstrese que el producto de dos operadores simétricos es simétrico si, y solo si, estos operadores son permutables.

2. Sea T el operador de un espacio E, que asocia a todo vector v su proyección ortogonal sobre un subespacio E' de E. Pruébese que T es simétrico. Recíprocamente, demuéstrese que todo operador simétrico tal que $T^2 = T$ (véase pág. 345) es una proyección ortogonal sobre un subespacio de E.

3. Pruébese que el núcleo y el subespacio imagen de un operador simétrico son dos subespacios ortogonales y suplementarios.

14. Función bilineal en un espacio euclidiano.—Sea una función bilineal f, definida en un espacio euclidiano E, que hace corresponder al par de vectores $(\mathbf{v}, \mathbf{w})$ el escalar real $f(\mathbf{v}, \mathbf{w})$. Vamos a demostrar que es posible asociar, de manera intrínseca, a la función f dos operadores lineales traspuestos uno del otro.

Sean $\mathbf{w}$ un vector dado de E y $\mathbf{v}$ arbitrario. La aplicación de E en R

$$\mathbf{v} \longrightarrow f(\mathbf{v}, \mathbf{w})$$

es lineal con relación a $\mathbf{v}$. Se le puede asociar (teorema de la pág. 440) un vector bien determinado $\mathbf{w}'$ tal que

$$\forall \mathbf{v}, \quad f(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot \mathbf{w}'.$$

Queda así definido un operador T: $\mathbf{w} \longrightarrow \mathbf{w}' = T\mathbf{w}$, cuya linealidad se deduce fácilmente de la de f respecto de $\mathbf{w}$.

Igualmente, a $\mathbf{v}$ dado se le puede hacer corresponder un $\mathbf{v}_1 = T_1\mathbf{v}$ tal que

$$\forall \mathbf{w}, \quad f(\mathbf{v}, \mathbf{w}) = \mathbf{v}_1 \cdot \mathbf{w}.$$

Los dos operadores T y T_1 así definidos son traspuestos uno del otro; en efecto, $\forall \mathbf{v}$, $\forall \mathbf{w}$, se tiene:

$$f(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot T\mathbf{w} = \mathbf{w} \cdot T_1\mathbf{v}.$$

Matrices.—En una base ortonormada, las matrices de estos operadores coinciden con las matrices asociadas a la forma; en efecto, sabemos que (véase pág. 426)

$$f(\mathbf{v}, \mathbf{w}) = {}^tXAY = {}^tY{}^tAX.$$

Haciendo $Y' = AY$, resulta $f(\mathbf{v}, \mathbf{w}) = {}^tXY'$. Suponiendo la base ortonormada, se trata de un producto escalar: el del vector $\mathbf{v}$ de matriz X por el vector de matriz Y'. Puesto que X es arbitrario, Y' es la matriz del vector $\mathbf{w}'$ anteriormente definido, por lo que A es la matriz de T en la base considerada. Se ve igualmente que ${}^tA$ es la matriz del traspuesto T_1.

15. Función bilineal y simétrica.—Supongamos ahora f bilineal y simétrica. Los dos operadores T y T_1 asociados a tal función coinci-

den. En efecto, la igualdad

$$f(\mathbf{v}, \mathbf{w}) = f(\mathbf{w}, \mathbf{v})$$

implica

$$\mathbf{v} \cdot T\mathbf{w} = \mathbf{w} \cdot T\mathbf{v}.$$

Recíprocamente, a todo operador simétrico T se le puede hacer corresponder una función bilineal y simétrica f, tal que

$$f(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot T\mathbf{w} = \mathbf{w} \cdot T\mathbf{v}.$$

En toda base ortonormada coinciden las matrices del operador y de la función.

Vectores propios.—Caractericemos directamente los vectores propios de T; sea $T\mathbf{v} = \lambda\mathbf{v}$ uno de estos vectores; para w arbitrario, tenemos:

$$f(\mathbf{v}, \mathbf{w}) = \mathbf{w} \cdot T\mathbf{v} = \lambda\mathbf{v} \cdot \mathbf{w}.$$

Por consiguiente, el valor de f para el vector propio v y para w arbitrario es proporcional al producto escalar $\mathbf{v} \cdot \mathbf{w}$, siendo el coeficiente de proporcionalidad, λ, valor propio. Recíprocamente, la condición

$$\mathbf{w} \cdot T\mathbf{v} = \lambda\mathbf{v} \cdot \mathbf{w}, \qquad \text{o} \qquad \mathbf{w} \cdot (T\mathbf{v} - \lambda\mathbf{v}) = 0$$

implica $T\mathbf{v} = \lambda\mathbf{v}$ puesto que w es arbitrario.

El teorema sobre la diagonalización demuestra que para toda función simétrica, f, es posible encontrar una base ortonormada al menos en la que la matriz de f es diagonal. En esta base, la forma cuadrática asociada $f_c(\mathbf{v})$ se convierte en

$$f_c(\mathbf{v}) = \Sigma\lambda_i(x^i)^2,$$

donde los λ_i son los valores propios de la matriz de f en una base cualquiera.

16. Diagonalización simultánea de dos funciones.—El estudio anterior conduce a determinar una base del espacio E en la cual f y el producto escalar admiten matrices diagonales. Vamos a completar este resultado con el fin de poner mejor en evidencia los papeles análogos que desempeñan estas dos funciones.

Se considera un espacio vectorial E en el que están definidas dos

funciones bilineales y simétricas f y g, de matrices A y B en una base de E:

$$f(\mathbf{v}, \mathbf{w}) = {}^t\mathrm{Y}\mathrm{A}\mathrm{X}, \qquad g(\mathbf{v}, \mathbf{w}) = {}^t\mathrm{Y}\mathrm{B}\mathrm{X}.$$

Vamos a demostrar que si una al menos de las formas cuadráticas asociadas es definida positiva, resulta posible encontrar una base de E (al menos) en la que las funciones f y g quedan diagonalizadas [1].

Supongamos, p. ej., $g_e(\mathbf{v})$ definida positiva; la función g puede desempeñar el papel de producto escalar en E, con lo cual este se convierte en un espacio euclidiano; estamos así en el caso anterior. En primer lugar, se determina en E una base ortonormada respecto de g, y sea A' la nueva matriz de f. Existe al menos una nueva base ortonormada en la que la matriz de f es diagonal, y cuyos elementos son los valores propios de A'; la matriz de g sigue siendo I. Se puede también hacer un estudio directo. A las funciones f y g (con g definida positiva) se asocia un operador lineal T tal que

$$f(\mathbf{v}, \mathbf{w}) = g(\mathbf{v}, \mathrm{T}\mathbf{w}) = g(\mathrm{T}\mathbf{v}, \mathbf{w}).$$

Este es un operador simétrico respecto de la función g, cuya matriz en la base inicial es, por otra parte, $\mathrm{B}^{-1}\mathrm{A}$. Estudiemos sus vectores propios: el $\mathbf{v}$ que corresponde al valor λ es tal que

$$(\forall \mathbf{w}) \qquad f(\mathbf{v}, \mathbf{w}) = \lambda g(\mathbf{v}, \mathbf{w})$$

lo que se escribe

$${}^t\mathrm{Y}\mathrm{A}\mathrm{X} = \lambda\, {}^t\mathrm{Y}\mathrm{B}\mathrm{X},$$

de donde

$$(\mathrm{A} - \lambda\mathrm{B})\mathrm{X} = 0$$

puesto que Y es arbitrario; ya que X no es nulo, se debe tener:

$$|\mathrm{A} - \lambda\mathrm{B}| = 0.$$

Los valores propios son, por tanto, las raíces de la ecuación anterior. La teoría de los operadores simétricos demuestra que estas son reales y que existe al menos una base formada por vectores propios, es decir, por vectores conjugados respecto de las dos formas cuadráticas.

---

[1] Esta condición suficiente no es, por otra parte, necesaria.

EJERCICIOS:

1. Sean dos funciones bilineales y simétricas de matrices A y B, y sean T y T' los operadores asociados. Se define una nueva función asociando a v el escalar $Tv \cdot T'v$. Exprésese la matriz de esta función; examínese el caso en el que A y B conmuten.

2. Diagonalícese en ejes ortonormados cada una de las formas cuadráticas siguientes del espacio R^3:

$$x^2-y^2+z^2+4xz; \qquad x^2-y^2+xz;$$
$$x^2+y^2+z^2-2yz-2zx+2xy; \qquad z^2-2xy.$$

Precísese la matriz de cambio.

3. OPERADOR ANTISIMÉTRICO.—El operador T del espacio euclidiano E se llama antisimétrico cuando ${}^tT=-T$.

a) Demuéstrese que T es antisimétrico si, y solo si,

$$\forall v, \forall w \qquad v \cdot Tw + w \cdot Tv = 0.$$

b) Pruébese que una condición necesaria y suficiente para que un operador sea antisimétrico es que transforme todo vector en un vector ortogonal.

c) ¿Cuál es el único valor propio que puede admitir un operador antisimétrico?

d) Si la dimensión de E es impar, demuéstrese que todo operador antisimétrico es singular (estúdiese el determinante de ${}^tT$).

e) Pruébese que el núcleo y el subespacio imagen T(E) de un operador antisimétrico T son ortogonales y suplementarios. ¿Qué cabe decir del operador inducido en T(E) por T? Dedúzcase de ello que el rango de un operador antisimétrico (o de una matriz antisimétrica) es un número par.

f) Caracterícense los operadores antisimétricos de un espacio de dos dimensiones.

g) Demuéstrese que, en un espacio de tres dimensiones, todo operador antisimétrico está caracterizado por un vector a, tal que $Tv = a \wedge v$ (producto vectorial).

h) Sean T antisimétrica y T_1 simétrica. Demuéstrese que la traza de TT_1 es nula.

ESPACIO UNITARIO DE *n* DIMENSIONES

La teoría anterior se aplica a un espacio vectorial real. El carácter definido positivo del producto escalar desempeña en ella un papel esencial [1]. Se obtiene una teoría análoga para un espacio vectorial com-

[1] Tenemos $f(v, v) \geq 0$ y $f(v, v) = 0$ implica $v = 0$. Una función bilineal definida en un espacio complejo no puede satisfacer condiciones análogas. En efecto, si el rango de f es n existirán bases en las que su matriz es I; tendremos así $f(v, v) = \Sigma(x^i)^2$, llamando x^i a las coordenadas de v. Sabemos que, en el cuerpo de los números complejos, la condición $\Sigma(x^i)^2 = 0$ no implica que sea $x^i = 0$.

plejo, modificando convenientemente la definición de bilinealidad. Resulta así el *espacio unitario*, cuya estudio engloba como caso particular el del espacio euclidiano.

17. Función hermítica.—Sea E un espacio vectorial de n dimensiones en el cuerpo de los números complejos. Se considera la aplicación h de $E \times E$ en C; al par $(\mathbf{v}, \mathbf{w})$ le corresponde el escalar $h(\mathbf{v}, \mathbf{w})$. Se dice que h es una *función hermítica* cuando satisface las propiedades siguientes:

a) Linealidad con relación al segundo argumento; es decir,

$$h(\mathbf{v}, \lambda\mathbf{w} + \mu\mathbf{w}') = \lambda h(\mathbf{v}, \mathbf{w}) + \mu h(\mathbf{v}, \mathbf{w}').$$

b) Antilinealidad con relación al primer argumento; es decir,

$$h(\lambda\mathbf{v} + \mu\mathbf{v}', \mathbf{w}) = \bar{\lambda} h(\mathbf{v}, \mathbf{w}) + \bar{\mu} h(\mathbf{v}', \mathbf{w})$$

donde $\bar{\lambda}$ y $\bar{\mu}$ designan los complejos conjugados de λ y μ.

Para $\lambda = \mu = 1$, se tienen las fórmulas:

$$f(\mathbf{v} + \mathbf{v}', \mathbf{w}) = f(\mathbf{v}, \mathbf{w}) + f(\mathbf{v}', \mathbf{w}).$$
$$f(\mathbf{v}, \mathbf{w} + \mathbf{w}') = f(\mathbf{v}, \mathbf{w}) + f(\mathbf{v}, \mathbf{w}').$$

Se dice que f es *biaditiva*.

c) Simetría hermítica; es decir,

$$h(\mathbf{w}, \mathbf{v}) = \overline{h(\mathbf{v}, \mathbf{w})}.$$

El cambio de orden de los vectores $\mathbf{v}$ y $\mathbf{w}$ transforma el valor de h en su conjugado. En particular, para $\mathbf{v} = \mathbf{w}$, $h(\mathbf{v}, \mathbf{v})$ es un número real, ya que es igual a su conjugado.

Representación matricial.— Sean $\{\mathbf{e}_i\}$ una base de E y $\mathbf{v} = \Sigma x^i \mathbf{e}_i$, $\mathbf{w} = \Sigma y^j \mathbf{e}_j$ dos vectores cualesquiera. Utilizando la linealidad y la antilinealidad, se obtiene la fórmula:

$$h(\Sigma x^i \mathbf{e}_i, \Sigma y^j \mathbf{e}_j) = \sum_{i,j} h(\mathbf{e}_i, \mathbf{e}_j) \bar{x}^i y^j. \qquad [1]$$

Hagamos de nuevo $a_{ij} = h(\mathbf{e}_i, \mathbf{e}_j)$ y consideremos la matriz $H = [a_{ij}]$ (i, índice de fila). La simetría hermítica entraña

$$a_{ji} = \bar{a}_{ij}.$$

Dos términos simétricos con relación a la diagonal principal son conjugados, lo que caracteriza a una matriz hermítica (véase página 311).

Si X e Y son las matrices columna de v y w, la fórmula [1] se escribe

$$h(v, w) = X^*HY \qquad [1\ bis]$$

designando por X^* la matriz asociada de X.

Recíprocamente, se define una función hermítica dando su matriz (hermítica) en una base de E.

Finalmente, en un cambio de base de matriz P la nueva matriz asociada a h se convierte en

$$H_1 = P^*HP.$$

El carácter hermítico se conserva.

OBSERVACIÓN.—Desarrollando una expresión tal como X^*HY, con H hermítica, obtenemos una *forma hermítica* en x^i e y^j. La forma es la expresión de la función en una base y depende de la elección de esta. Los dos términos se suelen emplear, sin embargo, indistintamente.

EJERCICIOS:

1. Empleando el cálculo matricial, compruébese el carácter hermítico de la función h *definida* por la fórmula [1 *bis*].

2. Demuéstrese que toda función f que verifica las propiedades *a)* y *b)* anteriores y tal que para v arbitrario $f(v, v)$ sea real, es una función hermítica. Desarróllese $f(v+w, v+w)$ haciendo uso de *a)* y *b)*.

3. Sea f una función que satisface dichas propiedades *a)* y *b)*. Defínase su matriz A en una base y demuéstrese que siempre se verifica $f(v, w) = X^*AY$.

Se hace corresponder a f su *asociada* f^*, tal que

$$f^*(v, w) = f(w, v).$$

Pruébese que f^* satisface igualmente *a)* y *b)*. Compárense las matrices de f y f^* en una base cualquiera.

4. ¿Cuál es la condición para que el producto de dos matrices hermíticas A y B sea hermítico?

18. Espacio unitario: producto escalar.—Se llama *espacio unitario* un espacio vectorial complejo dotado de una función hermítica h *definida positiva;* es decir, tal que

$$(\forall v), \quad h(v, v) \geqslant 0,$$

donde la condición $h(v, v) = 0$ implica que $v = 0$.

La función h se llama *producto escalar;* el escalar $h(\mathbf{v}, \mathbf{w})$ se denota corrientemente por $(\mathbf{v}, \mathbf{w})$. Satisface las propiedades *a), b)* y *c)* de la página 464. Conviene, pues, respetar el orden de los dos vectores del producto.

Se puede asociar a todo vector $\mathbf{v}$ un escalar real, positivo o nulo, su *norma:*

$$|\mathbf{v}| = \sqrt{(\mathbf{v}, \mathbf{v})}.$$

La condición $|\mathbf{v}| = 0$ implica $\mathbf{v} = \mathbf{0}$. Tenemos también:

$$|\lambda \mathbf{v}| = |\lambda| \cdot |\mathbf{v}| \qquad (|\lambda| \text{ módulo del número } \lambda).$$

También aquí es posible expresar el producto escalar con ayuda de las normas, pero la fórmula no es tan sencilla como en el espacio euclidiano. Poniendo $(\mathbf{v}, \mathbf{w}) = \alpha + i\beta$, se tiene:

$$|\mathbf{v} + \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 + (\mathbf{v}, \mathbf{w}) + (\mathbf{w}, \mathbf{v}) = |\mathbf{v}|^2 + |\mathbf{w}|^2 + 2\alpha,$$
$$|\mathbf{v} + i\mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 + i[(\mathbf{v}, \mathbf{w}) - (\mathbf{w}, \mathbf{v})] = |\mathbf{v}|^2 + |\mathbf{w}|^2 - 2\beta.$$

Se deduce fácilmente $\alpha + i\beta$ con ayuda de las normas de $\mathbf{v}$, $\mathbf{w}$, $\mathbf{v} + \mathbf{w}$, $\mathbf{v} + i\mathbf{w}$.

Determinación del producto escalar.—En la práctica se elige la matriz H en una base. No precisaremos las condiciones que es necesario imponer a H para asegurar el carácter definido positivo. Suponiéndolas realizadas, es siempre posible (véase más adelante) determinar bases ortonormadas en las que la función admite la matriz unidad I: para una de estas bases,

$$(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad (i \neq j) \qquad \text{y} \qquad (\mathbf{e}_i, \mathbf{e}_i) = 1.$$

Recíprocamente, la elección arbitraria de una base, supuesta ortonormada, define un espacio unitario. En efecto, la forma hermítica h así obtenida es tal que

$$h(\mathbf{v}, \mathbf{w}) = \mathrm{X}^* \mathrm{Y},$$

llamando X e Y a las matrices columna de $\mathbf{v}$ y $\mathbf{w}$ en la base considerada. Puesto que

$$h(\mathbf{v}, \mathbf{v}) = \mathrm{X}^* \mathrm{X} = \sum_1^n |x_i|^2,$$

la condición $h(\mathbf{v}, \mathbf{v}) = 0$ exige $|x_i| = 0$ (i de 1 a n), de donde $x_i = 0$ y $\mathbf{v} = \mathbf{0}$. Se trata, pues, de un producto escalar.

19. Ortogonalidad. Base ortonormada.—Dos vectores de producto escalar nulo se llaman *ortogonales*. La relación así definida es simétrica. El lector comprobará fácilmente que son válidas todas las consecuencias examinadas en la página 442 para el espacio euclidiano. Sin embargo, en los productos escalares conviene tener en cuenta el orden.

La desigualdad de Schwartz se escribe aquí

$$|\mathbf{v}, \mathbf{w}| \leqslant |\mathbf{v}| \cdot |\mathbf{w}|.$$

La igualdad supone $\mathbf{v}$ y $\mathbf{w}$ proporcionales.

Método de ortogonalización de Schmidt. TEOREMA.—*En un espacio unitario de n dimensiones todo subespacio de p dimensiones puede engendrarse por p vectores unitarios y ortogonales dos a dos.*

La demostración de la página 445 conserva su validez si se precisa el orden de los términos en los productos escalares que determinan los números λ, λ', μ', ... [1].

En particular, el teorema se aplica a todo el espacio E, de donde resulta la existencia de bases ortonormadas; en una base de este tipo,

$$(\mathbf{v}, \mathbf{w}) = \sum_{1}^{n} \bar{x}^i y^i = \mathrm{X}^* \mathrm{Y}.$$

Resulta igualmente que todo subespacio $\mathrm{E}' \subset \mathrm{E}$ admite un subespacio ortogonal y suplementario E'', que es único.

Consecuencia.—Vamos a demostrar, empleando la propiedad anterior, que un espacio unitario puede identificarse con su dual [2].

Sea $f \neq 0$ una función lineal definida en E. Sabemos (véase pág. 346) que el núcleo de f es un subespacio E' de dimensión $n-1$, que admite un subespacio ortogonal E'' de dimensión uno; sea $\mathbf{e}$ un vector unitario de E''. La función g tal que

$$\forall \mathbf{v}, \qquad g(\mathbf{v}) = f(\mathbf{v}) - \lambda(\mathbf{e}, \mathbf{v}) \qquad (\lambda \in \mathbb{C})$$

[1] Las fórmulas que dan λ' y μ' se escriben, p. ej.,

$$0 = (\mathbf{f}_1', \mathbf{f}_3) - \lambda' |\mathbf{f}_1'|^2, \qquad 0 = (\mathbf{f}_2', \mathbf{f}_3) - \mu' |\mathbf{f}_2'|^2.$$

[2] Esta demostración sirve también para el espacio euclidiano. Una demostración directa exigiría algunas precauciones debidas al carácter antilineal del producto escalar con relación al primer vector.

es lineal en v y nula para todo $v \in E'$. Elijamos λ de manera que

$$g(e) = f(e) - \lambda(e, e) = 0,$$

de donde

$$\lambda = \lambda_1 = f(e).$$

Para este valor particular λ_1, se tiene que $g(v)$ es idénticamente nula, de modo que

$$(\forall v) \quad f(v) = (\lambda_1 e, v) = (a, v) \quad (a = \lambda_1 e).$$

A toda función lineal corresponde un vector de E, y, recíprocamente, todo $a \in E$ define una función lineal. La correspondencia es lineal.

TEOREMA.—*Un espacio unitario es naturalmente isomorfo a su dual.*

20. Cambio de base. Matriz unitaria.—Caractericemos la matriz U de un cambio de base que asocia dos matrices ortogonales. Siendo iguales a I las matrices del producto escalar en ambas bases, la fórmula de la página 465 se convierte, con $P = U$, en

$$I = U^* I U = U^* U.$$

Por consiguiente, la asociada de U es también su inversa: $U^{-1} = U^*$, de donde $UU^* = I$. Esta condición es, además, suficiente para que la segunda base sea ortonormada si lo es la primera. Una matriz U que satisface la condición anterior se llama unitaria. Las matrices ortogonales aparecen como caso particular: son las matrices unitarias reales.

Los elementos de una matriz unitaria satisfacen las n^2 relaciones que resultan de $U^* U = I$ y las n^2 relaciones equivalentes que se deducen de $UU^* = I$. Poniendo

$$U = \begin{bmatrix} \alpha_1^1 & \alpha_2^1 & \dots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \vdots & & & \\ \alpha_1^n & \alpha_2^n & \dots & \alpha_n^n \end{bmatrix},$$

de donde

$$U^* = \begin{bmatrix} \bar{\alpha}_1^1 & \bar{\alpha}_1^2 & \dots & \bar{\alpha}_1^n \\ \bar{\alpha}_2^1 & \bar{\alpha}_2^2 & \dots & \bar{\alpha}_2^n \\ \vdots & \vdots & & \vdots \\ \bar{\alpha}_n^1 & \bar{\alpha}_n^2 & \dots & \bar{\alpha}_n^n \end{bmatrix}$$

se obtienen, bien las n^2 relaciones

$$\sum_{k=1}^{k=n} \bar{\alpha}_i^k \alpha_j^k = \delta_{ij} \qquad (U^* U = I)$$

que expresan que los vectores columna de la matriz U son normados y ortogonales dos a dos, o bien las n^2 relaciones

$$\sum_{k=1}^{k=n} \alpha_k^i \bar{\alpha}_k^j = \delta_{ij} \qquad (U U^* = I)$$

que expresan las mismas condiciones para los vectores columna de U^*, los cuales no son en este caso los vectores fila de U.

Consideremos, finalmente, el determinante de una matriz unitaria. Se tiene:

$$\text{det. } U \cdot \text{det. } U^* = 1.$$

Puesto que det. U^* es conjugado de det. U, resulta que estos dos números complejos tienen por módulo 1.

21. Asociado de un operador lineal. Grupo unitario.—La noción de operador asociado generaliza la de operador traspuesto del espacio euclidiano. Enunciemos el resultado.

TEOREMA.—*A todo operador lineal* T *de un espacio unitario se le puede hacer corresponder un operador lineal asociado* T^* *tal que para* v *y* w *arbitrarios se tenga*

$$(\mathbf{v}, T\mathbf{w}) = (T^*\mathbf{v}, \mathbf{w}).$$

En toda base ortonormada las matrices T y T^* son asociadas.

La demostración es análoga a la de la página 452. Partiendo de **a**, vector dado, se considera $(\mathbf{a}, T\mathbf{v})$ que es lineal en v y que, por consiguiente (véase pág. 468), puede escribirse en forma de producto escalar:

$$(\mathbf{a}, T\mathbf{v}) = (\mathbf{a}', \mathbf{v}). \qquad [1]$$

De ello se deduce una aplicación T^* tal que $\mathbf{a}' = T^*\mathbf{a}$ $(\forall \mathbf{a})$; demostremos que T^* es lineal. En primer lugar, es aditiva (propiedad del producto escalar) y, por otra parte, reemplazando **a** por $\lambda\mathbf{a}$ se multiplica $(\mathbf{a}; T\mathbf{v})$ por $\bar{\lambda}$, lo que cambia $\mathbf{a}'$ en $\lambda\mathbf{a}'$ en virtud de la antilinealidad

del producto escalar; de ello resulta la linealidad de T*. La igualdad [1] puede, pues, escribirse [1]:

$$(\forall v, \forall w) \quad (v, Tw) = (T^*v, w). \qquad [2]$$

Sean A y B las matrices en una base ortonormada; se obtiene:

$$(\forall X, \forall Y) \quad X^*AY = (BX)^*Y = X^*B^*Y$$

y, por tanto (véase pág. 452), $A = B^*$.

El lector comprobará que se generalizan las propiedades de los núcleos y de los subespacios imagen. Los valores propios de uno de los operadores son los conjugados de los valores propios del otro. Finalmente, dos subespacios propios, uno de T y el otro de T*, correspondientes a dos valores propios no conjugados, son ortogonales

Transformación unitaria.—Se llama transformación unitaria todo operador lineal U que conserva el producto escalar de dos vectores cualesquiera:

$$(v, w) = (Uv, Uw).$$

Se conserva la norma de todo vector y recíprocamente (según la relación existente entre norma y producto escalar estudiada en la pág. 466).

La transformación producto de dos transformaciones unitarias es unitaria, así como la inversa: el conjunto constituye el grupo unitario.

Por último, una condición necesaria y suficiente para que una transformación sea unitaria es que su inversa sea su asociada.

En toda base ortonormada la matriz de una transformación unitaria es unitaria. Todas estas propiedades se demuestran como las propiedades análogas del grupo ortogonal.

22. Operador hermítico.—Un operador lineal T idéntico a su asociado se llama *hermítico*: $T^* = T$. Para ello es necesario y suficiente que

$$\forall v, \forall w, \quad (v, Tw) = (Tv, w). \qquad [1]$$

Resulta de lo anterior que (v, Tv) es un número real. Recíprocamente, esta última condición es suficiente. Para verlo basta desarro-

[1] En el caso euclidiano, escribimos $v \cdot Tw = w \cdot Tv$ (o $Tv \cdot w$) sin tener en cuenta el orden de los vectores en un producto escalar. Aquí, para respetar la linealidad y la antilinealidad, ponemos v y T*v como primeros elementos, y Tw y w como segundos.

llar el producto escalar $(v+w, Tv+Tw)$; de ese modo se demuestra que $(v, Tw)+(w, Tv)$ debe ser real; es decir, que

$$(v, Tw)=\overline{(w, Tv)}=(Tv, w).$$

Esta es la condición [1]. En toda base ortonormada la matriz de un operador de este tipo es hermítica.

Finalmente, la condición [1] puede aún reemplazarse por la siguiente: para que T sea hermítica es necesario y suficiente que la función h definida por

$$h(v, w)=(v, Tw) \qquad [2]$$

sea hermítica. Recíprocamente, toda función hermítica h define un operador hermítico T tal que se verifica la igualdad [2][1].

Diagonalización. TEOREMA 1.—*Los n valores propios de un operador hermítico son reales. Los subespacios propios correspondientes a dos valores propios distintos son ortogonales.*

Sea λ un valor propio y v un vector propio asociado; puesto que $Tv=\lambda v$, se tiene:

$$(v, Tv)=(v, \lambda v)=\lambda(v, v).$$

Hemos visto que (v, Tv) es real; puesto que también lo es (v, v), resulta igualmente que lo es λ. Este lenguaje geométrico interpreta el cálculo efectuado en la página 456 en el caso de un espacio euclidiano.

Por último, la ortogonalidad de dos subespacios propios se demuestra como en la página 458.

TEOREMA 2.—*Existe al menos una base ortonormada de un espacio unitario en la que un operador hermítico admite una matriz diagonal. Los elementos diagonales son los valores propios.*

El mismo método que para un operador simétrico (véase pág. 458). Si e_1 es un vector propio que corresponde al valor propio λ_1, el subespacio ortogonal a e_1 es invariante. El operador inducido, que es también hermítico, admite un vector propio...

---

[1] La misma demostración que en el espacio euclidiano (véase pág. 460) haciendo uso de la propiedad de la página 468.

EJERCICIOS:

1. Demuéstrese que todo valor propio de una transformación unitaria tiene por módulo 1. Caso de las matrices ortogonales (véase ejercicio 1, página 459).

2. Compruébese directamente que los coeficientes del polinomio característico de una matriz hermítica son reales.

3. Sea A una matriz antisimétrica real, ¿qué se puede decir de iA? Demuéstrese que existe una matriz unitaria U tal que $U^{-1}AU$ es diagonal, siendo los elementos diagonales imaginarios puros.

Dedúzcase de ello que una matriz antisimétrica real de orden impar es singular.

4. Un operador T del espacio unitario se llama *normal* cuando es permutable con su asociado: $TT^* = T^*T$. Compruébese que los operadores unitarios y los operadores hermíticos son normales. Demuéstrese que una condición necesaria y suficiente para que un operador sea normal es que

$$\forall v, \forall w, \quad (Tv, Tw) = (T^*v, T^*w).$$

Examínese el caso $v = w$. Dedúzcase de ello que los núcleos de T y de T^* se confunden. Pruébese que el núcleo y la imagen de un operador normal son ortogonales y suplementarios.

5. Demuéstrese que si T es normal todo vector propio de T es vector propio de T^*, siendo conjugados los valores propios. (Efectúese la composición de los operadores $T - \lambda I$ y $T^* - \lambda I$.)

6. Pruébese que todo operador normal de un espacio unitario posee al menos un sistema de vectores propios ortogonales dos a dos. Si λ_i es el valor propio de T y e_i un vector propio asociado, demuéstrese que el subespacio ortogonal a e_i es invariante respecto a T (hágase uso del ejercicio anterior). Estúdiese el operador inducido por T en el subespacio y dedúzcase de ello (por recurrencia) la propiedad pedida.

EJERCICIOS DE RECAPITULACION

1. Con relación a dos ejes rectangulares se considera la transformación

$$X = x, \qquad Y = \frac{3}{2}x + y,$$

de la que llamamos A a la matriz

1.º Calcúlese la traspuesta A' y los productos AA' y A'A.

2.º Demuéstrese que $AA' = B$ es de la forma $P^{-1}CP$, donde C es una matriz diagonal. Determínense P y C.

3.º Demuéstrese que existe D tal que $D^2 = C$; calcúlese D.

4.º Pruébese que existe H simétrica tal que $H^2 = B$. Calcúlese H.

5.º Demuéstrese que $A = HU$ con U ortogonal. Calcúlese U.

2. Se considera el espacio euclidiano R^n y el subespacio R^m engendrado por los m primeros vectores de la base canónica. Hállese el coseno del menor ángulo formado por el vector $(1, 1, \ldots, 1) \in R^n$ con el subespacio anterior.

Resp.: $\sqrt{\frac{m}{n}}$

3. Demuéstrese que en un espacio euclidiano de n dimensiones no se puede hallar un conjunto de $n+2$ vectores que formen ángulos obtusos dos a dos.

Indicaciones: Los $n+1$ primeros vectores $v_1, v_2, \ldots, v_{n+1}$ son dependientes:

$$\sum_{1}^{n+1} \lambda_i v_i = 0.$$

Multiplicando escalarmente por v_{n+2} se deduce que los coeficientes λ_i no son todos del mismo signo. Hacemos $w_1 = \Sigma \lambda_i v_i$ (con $\lambda_i > 0$) y $w_2 = \Sigma \lambda_i v_i$ (con $\lambda_i < 0$) y se demuestra que la igualdad $w_1 = -w_2$ (de donde $w_1^2 = -w_1 \cdot w_2$) conduce a una imposibilidad.

4. Resuélvase y discútase la ecuación matricial

$$\begin{bmatrix} m & 6 \\ 2 & 12 \end{bmatrix} \begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} m+4 & 40 \\ 10 & 80 \end{bmatrix},$$

en la que m es un parámetro.

a) En el caso particular en el que la solución no es única las soluciones forman una familia (F).

¿Existe en (F) una matriz que corresponda geométricamente a una homotecia?

b) ¿Existe en (F) una matriz que corresponda en ejes ortonormados a una rotación de $\frac{2\pi}{3}$? Calcúlense sus elementos.

c) ¿Existe en (F) una subfamilia (*f*) para la que la ecuación característica tenga una raíz doble? Estúdiese (*f*) desde el punto de vista geométrico.

5. Sea T un operador de R^n del que los valores propios se suponen reales pero no necesariamente distintos. Demuéstrese que es siempre posible hallar al menos una base ortonormada en la que la matriz de T es triangular.

Indicaciones: Si e_1 es un vector propio de T asociado a λ_1, se define en el subespacio E' ortogonal a e_1 un operador T' tal que, si $v \in E'$, sea T'v la proyección ortogonal de Tv sobre E'. Estúdiese T': ¿cuáles son sus valores propios?

Ejemplos: *a)* Aplíquese a T de matriz $A = \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}$ en una base ortonormada.

b) La misma cuestión para T_1, de matriz

$$A_1=\begin{bmatrix} -2 & 1 & 3 \\ -3 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix}.$$

Resp.: Cabe tomar, respectivamente, como matrices de cambio

$$P=\frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}; \qquad P_1=-\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \end{bmatrix}.$$

6. Se considera la matriz

$$S=\begin{bmatrix} \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{2}{3} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{2} \end{bmatrix}$$

en la que la suma de los términos de cada columna es igual a la unidad.

1.º Calcúlense los valores propios de S.

2.º Demuéstrese que, para el mayor valor propio, existe un vector propio (x, y, z) tal que

$$0\leqslant x\leqslant 1, \qquad 0\leqslant y\leqslant 1, \qquad 0\leqslant z\leqslant 1, \qquad x+y+z=1.$$

3.º Hállese una matriz diagonal B semejante a S y calcúlese una matriz de cambio P; obténgase P^{-1}.

4.º Determínese el límite de B^n cuando n tiende a infinito y dedúzcase del mismo el de S^n. ¿Qué observación cabe hacer sobre los vectores columna de esta matriz?

5.º Considérense de nuevo estas cuestiones (o sus análogas) con la matriz

$$S=\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$$
$$0\leqslant a\leqslant 1, \qquad 0\leqslant a\leqslant 1.$$

Tales matrices se llaman *estocásticas* e intervienen en el cálculo de probabilidades.

7. Se consideran las seis *matrices de base*

$$E=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad B=\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

$$Q=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad R=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad S=\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y la familia $\mathfrak{F}$ de matrices de la forma

$$xE+yA+zB+uQ+vR+wS,$$

donde los coeficientes x, y, z, u, v, w pertenecen a Z. Una matriz que pertenezca a $\mathfrak{F}$ y admita una inversa perteneciente también a $\mathfrak{F}$ se llama $\mathfrak{F}$-invertible.

1.º Demuéstrese que las seis matrices de la base constituyen un grupo para la multiplicación. Indíquese un grupo isomorfo. ¿Son $\mathfrak{F}$-invertibles las matrices de la base?

2.º Fórmese la matriz

$$U=-A+B-Q+R+S$$

y calcúlese su cuadrado.

Compruébese el resultado calculándolo con ayuda de la tabla de multiplicar construida en el apartado 1.º.

¿Es U $\mathfrak{F}$-invertible?

¿Es posible predecir sus valores propios sin efectuar ningún cálculo? Calcúlense sus valores propios.

¿Es semejante U a una matriz diagonal?

3.º Sea ${}^tU$ la traspuesta de U.

Demuéstrese que $V=U{}^tU$ y $V_1={}^tUU$ son $\mathfrak{F}$-invertibles.

4.º Se considera la matriz

$$W=E+A-B+R-S.$$

Hállese la ley de formación de W^n y de W^{-n}.

¿Se pueden formar matrices $\mathfrak{F}$-invertibles a partir de W?

INDICE ALFABETICO

INDICE ALFABETICO
DE AUTORES Y MATERIAS

ABEL, 54.
Absorción, ley de, 40.
Algebra:
  de Boole, 42.
  de los conjuntos, 50.
  del grupo, 204.
  noción de, 41.
  o sistema hipercomplejo, 204.
Algoritmo de Euclides, 235, 287.
Altura de una matriz, 291.
Angulo de dos vectores, 444.
Anillo:
  booleano, 114.
  cociente, 95.
  compuesto, 97.
  conmutativo, 78.
  de cuadrado nulo, 79.
  de factorización única, 280.
  de integridad, 78.
  de las clases residuales, 86, 246.
  de las matrices cuadradas de orden *n*, 317.
  de los endomorfismos de un espacio vectorial, 343.
  de los enteros relativos, 82.
  de los polinomios:
    sobre un anillo, 215.
    sobre un cuerpo, 218.
  estructura de, 75.
  euclidiano, 287.
  ordenado, 79.
Anillos, compuesto directo de, 113.
Antecedente de un elemento, 7.
Aplicación:
  alternada, 369.
  antisimétrica, 368.
  asociada a una función, 426.
  bilineal, 425.
  biyectiva, 9.
  idéntica, 8.
  inyectiva, 9.
  lineal, 334.
  multilineal, 364.
  noción de, 7.
  regular, 339.
  suprayectiva, 8.
  traspuesta, 360.
ARGAND, 188.
Argand-Cauchy, plano de, 198.
Argumento de un número complejo, 194.
Arquímedes, axioma de, 180.
Asociado de un operador lineal, 469.
Automorfismo, 38, 135.
  del cuerpo de los números complejos, 188.
  interno, 69.

Base:
  cambio de, 136, 347, 358, 426, 429, 448, 468.
  canónica, 135.
    de las matrices (m, n), 296.
  de la función logarítmica, 172.
  de un espacio:
    de funciones lineales, 346.
    vectorial, 132.
  de un grupo abeliano, 74.
  dual, 350, 356, 426.
  natural, 134.
  ortonormada, 445, 467.
Bezout, teorema de, 234.
Biyección:
  noción de, 9.
  recíproca, 9.
  regular, 36.
BOMBELLI, 188.
Boole, álgebra de, 42.

Cajas:
  operaciones con, 315.
  subdivisión de una matriz en, 293.
Cálculo de un determinante, 383.
CANTOR, 3, 151.

Cantor-Dedekind, axioma de, 162, 180.
Característica:
de un anillo, 90.
de un cuerpo, 103.
Cardinal, número, 20.
CAUCHY, 151.
criterio de, 151.
Cayley:
problema de, 53.
teorema de, 62.
Clase:
asociada a una propiedad, 11.
de asociación para la divisibilidad, 93.
de congruencia:
de las matrices simétricas, 429.
para los polinomios, 222.
de equivalencia, 16.
de semejanza de una matriz, 409.
residual:
en un anillo de polinomios, 246.
en Z, 70, 72, 86.
Clases de matrices, 329.
Codimensión de un subespacio vectorial, 146.
Cofactor, 380.
Combinación:
con repetición, 48.
lineal, 118.
simple, 46.
Complejo en un grupo, 65.
Complementación, noción de, 41.
Compuesta de dos aplicaciones lineales, 342.
Compuesto directo:
de dos anillos, 113.
de dos grupos, 73.
Conjunto:
casi ordenado, 19.
cociente, 16.
de las partes, 5.
finito, 3.
infinito, 3.
numerable, 20.
ordenado, 18.
perfecto de medida nula, 182.
producto, 5.
triádico de Cantor, 182.
vacío, 5.
Conjuntos equipotentes, 9.
Conmutador, noción de, 56.
Convergencia sobre un grupo ordenado, 110.
Construcción de Cauchy para los complejos, 188.
Coordenadas:
contravariantes, 359.
covariantes, 359.
Coseno, relación del, 444.
Criterio:
de Cauchy, 151.
de Eisenstein, 285.
Cuadrado latino, 59.
Cuantificador, noción de, 11.
Cuaternio, noción de, 202.
Cuerpo:
de las fracciones racionales, 262.
de los complejos:
por clases, 247.
por pares, 184.
de los racionales, 104.
de los reales, 151.
finito, 101.
noción de, 99.
ordenado, 108.

D'Alembert, teorema de, 251.
Dependencia lineal, 127.
Derivada de un polinomio, 257.
Derivadas sucesivas, 259.
Desarrollo de un determinante, 380.
Descomposición:
de las fracciones racionales en elementos simples, 263.
de un polinomio en factores primos, 242.
de una forma cuadrática en cuadrados, 432.
sobre C, 251.
sobre R, 251.
Desigualdad:
de Cauchy-Schwartz, 446.
de Minkowski, 443.
de Schwartz, 436, 442, 467.
Determinación de una base para un subespacio vectorial, 140.
Determinante:
de un endomorfismo, 406.
de un producto de matrices, 378.
de una matriz, 372, 375.

Diagonalización:
de un operador:
hermítico, 471.
lineal, 411.
simétrico, 458.
de una forma, 432.
simultánea de dos funciones, 44.
Diagrama:
de una relación de orden, 18.
de Venn, 40.
Dimensión:
de un espacio vectorial, 132.
de un subespacio, 137.
Dimensiones de una matriz, 291.
Divisibilidad:
en los dominios de integridad, 91.
por $(x-a)$, 227.
División:
euclidiana, 222.
por $(x-a)$, 226.
según las potencias crecientes, 229.
Divisor:
de cero, 78, 89, 318.
estricto, 93.
primo, 93.
Dominio:
de integridad, 78, 90.
de operadores para una ley, 22.

Ecuación:
de cuarto grado, 255.
de segundo grado, 192.
de tercer grado, 254.
Ecuaciones lineales, 139, 302.
Elemento:
de un conjunto, 3.
inverso, 30.
neutro, 29.
opuesto, 30.
positivo de un anillo ordenado, 79.
primo, 93.
regular:
a la derecha, 30.
a la izquierda, 30.
simétrico, 30.
unidad de un anillo, 76.
Elementos:
asociados para la divisibilidad, 92.
conjugados en un grupo, 67.
equivalentes, 16.
unitarios de un anillo, 96.
Eliminación, nociones sobre la, 237.
Endomorfismos de un espacio vectorial, 343, 405.
Enteros:
de Gauss, 288.
naturales, 36.
relativos, 83, 90.
Equivalencia:
asociada a una aplicación, 16.
regular sobre un grupo, 69.
Espacio vectorial:
de las aplicaciones de E en F, 341.
de matrices rectangulares, 295.
de n dimensiones, 132.
de polinomios sobre un cuerpo, 214.
dual, 354, 440, 468.
euclidiano, 439.
noción de, 115.
unitario, 463.
Especialización, principio de, 220.
Extensión:
algebraica de un cuerpo, 246.
de una ley, 23.

Factorización única, 280.
Forma:
bilineal, 313, 426.
cuadrática, 314, 428.
definida positiva, 436.
fundamental en un espacio euclidiano, 440.
hermítica, 465.
indefinida, 437.
lineal, 346.
polar, 430.
semidefinida, 437.
simétrica, 456.
Formas cuadráticas sobre R, 414
Fórmula de Taylor, 259.
Fórmulas:
de Cardano, 255.
de Moivre, 196.
de Newton, 285.
sumatorias, 51.
Función:
aditiva de conjunto, 50.
aplicación y, 7.
bilineal:

en un espacio euclidiano, 460.
y simétrica, 460.
continua, 163.
determinante, 372, 450.
exponencial, 172.
hermítica, 464.
logarítmica, 172.
p veces lineal, 366.
racional entera, 366.
Funciones simétricas fundamentales, 275.

GALOIS, 54.
campo de, 251.
teorema de, 74.
Gauss:
enteros de, 288.
método de, 433.
Grado de un polinomio, 212.
Gráfica:
de una aplicación, 8.
de una relación binaria, 13.
Grupo:
abeliano, 54.
aditivo:
de C, 198.
de R, 154, 168.
cíclico, 72.
de las matrices regulares, 323.
de las rotaciones en el espacio, 207.
de sustituciones, 60.
estructura de, 54.
factor, 70.
lineal, 136, 407.
monógeno, 72.
multiplicativo:
de C, 198.
de R, 155.
ordenado, 112.
ortogonal, 454.
simétrico, 61.
unitario, 469.
Grupos finitos, 57, 59.

Hamilton, cuaternios de, 202.
Homomorfismo, noción de, 38.
Horner, esquema de, 241.
Ideal:
bilateral, 94.
de un anillo de polinomios, 231.
Imagen:
de un elemento, 7.
de un subespacio, 337.
de una parte, 8.
homomorfa:
de un anillo, 94.
de un conjunto, 38.
de un grupo, 71.
recíproca:
de un subespacio de F, 338.
de una parte, 8.
Independencia:
criterio de, 129.
en un espacio euclidiano, 442
lineal, 127.
teorema fundamental de, 130.
Indice:
de un subgrupo finito, 66
libre, 119.
mudo, 119.
Intersección:
de dos conjuntos, 5, 40.
de subespacios vectoriales, 122, 138.
Intervalos, principio de los, 161.
Inversa de una matriz cuadrada, 321, 326, 387.
Inyección, noción de, 9.
Isomorfismo:
de espacios vectoriales, 120, 134.
de grupos, 55.
entre matrices y operadores, 403.
extensión de un teorema de, 174, 178.
fundamental en R, 169.
natural, 135.
noción de, 35.

Kronecker:
condición de, 303.
símbolo de, 137.
teorema de, 302.

Ley:
asociativa, 25.
conmutativa, 27.
de composición, 22.
de Sylvester, 434.
definida en todo el conjunto, 22.
distributiva respecto de otra, 28.
estable respecto de una equivalencia, 24.

externa, 22.
inducida sobre un conjunto cociente, 28.
interna, 22.
producto, 24.
Límite:
inferior, 19, 162.
superior, 19, 162.
Longitud:
de una matriz, 291.
en un espacio euclidiano, 443.

Matrices:
elementales, 314.
monomias, 333.
ortogonales, 448.
semejantes, 409.
Matriz:
adjunta, 387.
antisimétrica, 296.
asociada, 34.
columna, 293.
completa de un sistema, 144.
conjugada, 311.
de cambio, 324, 359.
de las coordenadas de un sistema de vectores, 141.
de los coeficientes, 144, 145.
de permutación, 315.
de sustitución, 315.
de una aplicación lineal, 401.
de una relación, 13.
de verdad, 293.
definición general de, 291.
diagonal, 297, 320.
escalar, 320.
fila, 293.
hermítica, 34.
reducida, 142.
o canónica de una clase de semejanza, 411.
regular, 301.
regularmente escalonada, 142.
simétrica, 295.
singular, 301.
triangular, 143, 297.
unitaria, 468.
Mayorante de un subconjunto de un conjunto ordenado, 19.
M. C. D. de polinomios, 232, 244
M. C. M. de polinomios, 236, 244.
Menor de un elemento de un determinante, 380.
Minkowski, desigualdad de, 443.
Minorante de un subconjunto de un conjunto ordenado, 19.
Módulo de un número complejo. 189, 194.

Norma:
de un cuaternio, 203.
de un vector, 443, 466.
en ciertos anillos. 96.
Núcleo:
de un homomorfismo de grupos, 71.
de una aplicación lineal, 339. 344, 362.
de una función bilineal, 429.
Números:
complejos, 184.
enteros:
naturales, 36.
relativos, 82.
p-ádicos, 167.
racionales, 104.
reales, 151.

Operaciones:
con cajas, 315.
con las funciones lineales, 348.
elementales con los sistemas de vectores, 123.
Operador:
diagonalización de un, 411.
en un espacio euclidiano, 452.
hermítico, 470.
ortogonal, 452.
simétrico, 456.
singular, 407.
traspuesto, 452.
Operadores lineales, 399.
Orden:
de infinitud, 19.
de un grupo finito, 57.
estricto, 18.
parcial, 18.
total, 18.
Ordinal, noción de, 20.
Orientación de un espacio vectorial, 374.

Ortogonales, vectores, en el espacio:
euclidiano, 442.
unitario, 467.
Ortogonalización, método de Schmidt, 445, 467.

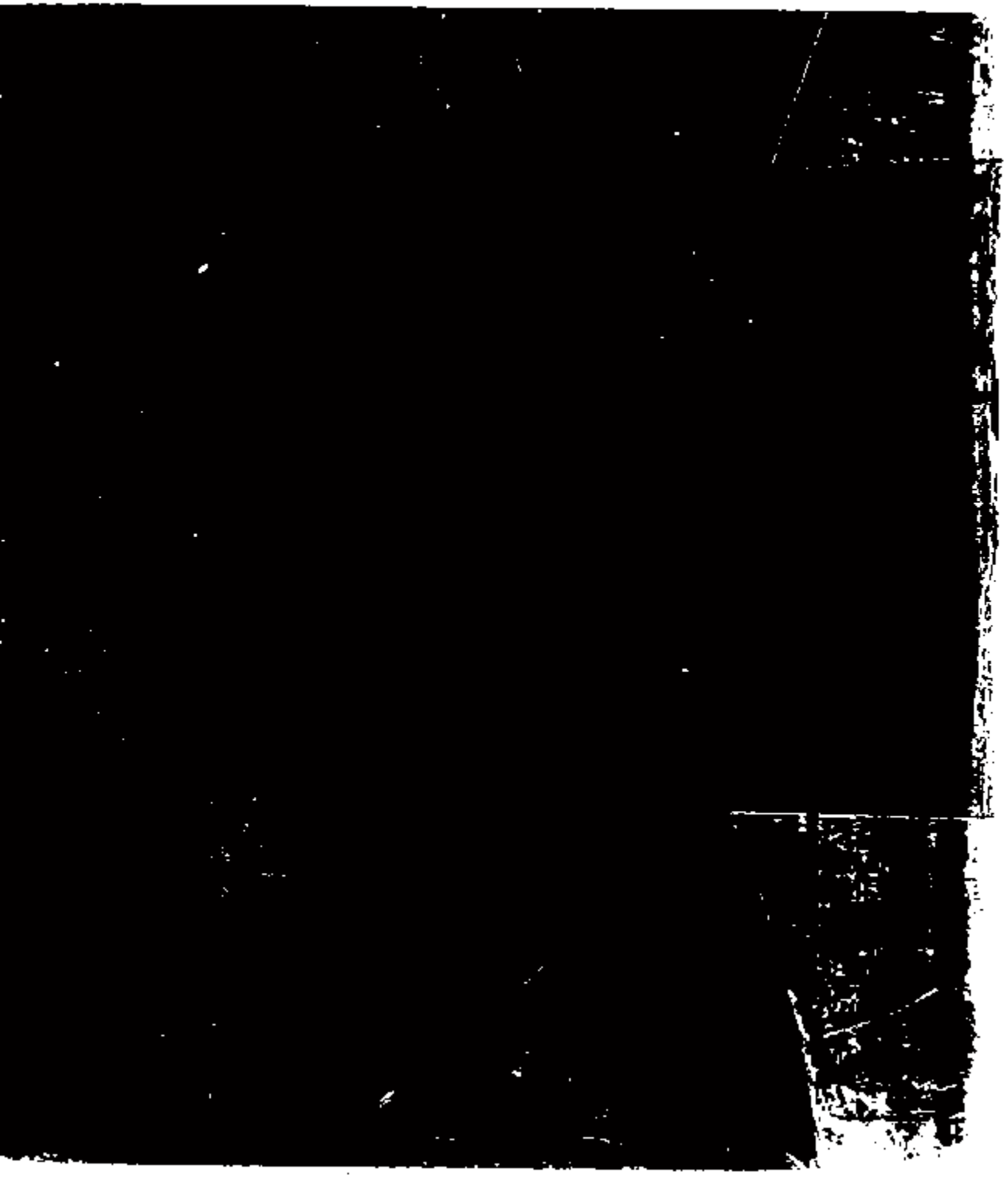
Paridad de una sustitución, 63.
Parte cerrada para una ley, 23.
Partición de un conjunto, 5.
Permutación:
con repetición, 47.
noción de, 45.
Plano de Argand-Cauchy, 198.
Polinomio:
característico, 414.
continuidad sobre R de un, 164.
formal, 211, 271.
irreducible, 242.
primitivo, 281.
Polinomios:
de varias variables, 27.
simétricos, 275.
Producto:
de conjuntos, 5.
de dos aplicaciones, 33.
de espacios vectoriales, 118.
de matrices, 305, 403.
de números complejos, 184.
escalar, en un espacio:
euclidiano, 439.
unitario, 466.
exterior, 391.
vectorial, 450.
Proyección:
paralelamente a un subespacio, 125.
ortogonal, 442.
P. T. O. L., noción de, 273.

Raíces:
múltiples de un polinomio, 260.
n-ésimas de la unidad, 200.
Raíz:
cuadrada de un número complejo, 191.
n-ésima de un número real, 164.
Rango:
de un sistema de vectores, 133.
de una aplicación, 338, 344, 362, 426.
de una función bilineal, 428.
de una matriz, 300, 329.
Referencial:
álgebra de las partes de un, 40.
noción de, 3.
Relación:
antisimétrica, 14.
binaria, 11.
circular, 17.
de divisibilidad, 18, 91.
de equivalencia, 15.
de orden, 18.
de recurrencia, 147.
del coseno, 444.
reflexiva, 13.
simétrica, 14.
transitiva, 14.
Relaciones entre los coeficientes y las raíces de una ecuación, 245.
Restricción:
de una aplicación, 8.
del cuerpo de escalares, 122.
Resultante de dos polinomios, 238.
Retículo, estructura de, 44.

Schmidt, método de, 445, 467.
Schwartz, desigualdad de, 436, 442, 467.
Signatura:
de una función bilineal, 435.
de una sustitución, 65.
Símbolo de Kronecker, 137.
Simetrización, teorema de, 98.
Sistema:
de generadores de un espacio vectorial, 123.
en cascada, 145.
libre, 128.
ligado, 127.
Sistemas:
de Cramer, 303, 388.
de ecuaciones lineales, 143, 301, 388.
equivalentes:
de ecuaciones, 144.
de vectores, 120.
Stone, teorema de, 42.
Subanillo:
de polinomios con relación a una matriz, 320.
noción de, 78.

Subconjunto:
amplio, 4.
complementario, 5.
disjunto, 4.
estricto, 4.
imagen de una aplicación, 8.
Subcuerpo, noción de, 103.
Subespacio:
engendrado por:
un sistema de vectores, 123.
una parte, 124.
fila, 329.
invariante, 410.
ortogonal suplementario, 447.
vectorial, 121.
Subespacios:
conjugados, 351.
propios, 458.
suplementarios, 125.
vectoriales disjuntos, 122.
Subgrupo:
distinguido, 66.
invariante, 66.
noción de, 59.
Subgrupos conjugados, 67.
Sucesión:
equivalente, 152.
a cero, 152.
negativa, 158.
opuesta, 155.
positiva, 158.
regular, 152.
inversa de una, 156.
Suma directa:
de grupos abelianos, 73.
de matrices, 411.
de subgrupos vectoriales, 122, 139, 412.
Suprayección, noción de, 8.
Sustitución:
lineal, 292, 312.
noción de, 60.
teorema de la, 131.
Sylvester, ley de, 435.

Tabla de composición, 23.
Teorema:
de Gauss para los polinomios, 234.
de homomorfismo:
para los grupos, 70.
para un conjunto dotado de una ley, 39.
de la base incompleta, 138.
Tipo:
de buen orden, 21.
de orden, 21.
Transmutado de un elemento de un grupo, 67.
Trasposición:
de matrices, 294.
de un producto de matrices, 310.
Traza de una matriz, 15.

Unión:
de conjuntos, 40.
finitos, 50.
de dos conjuntos, 40.

Valor:
absoluto en un anillo ordenado, 80.
propio, 411, 456.
Variable:
booleana, 42.
en general, 7.
Variación:
con repetición, 48.
noción de, 45.
Vector:
columna, 294.
fila, 294.
propio, 411, 458.
Vectores:
conjugados, 431.
ortogonales, 442, 467.